BRIDGE-MO: RISK-AWARE WITNESS-GUIDED OPTIMIZATION FOR MULTI-OBJECTIVE JOINT TARGET ATTAINMENT

Anonymous authors
Paper under double-blind review

ABSTRACT

Pareto-based multi-objective optimization is widely used in scientific design to balance competing objectives. However, Pareto optimality is often poorly matched to joint target attainment, where success requires all properties to cross fixed thresholds simultaneously. Pareto methods may continue improving already-satisfied objectives or expanding the trade-off frontier while a bottleneck property remains below target. To address this issue, we introduce Bayesian Risk-Aware Intervention-Driven Guided Exploration for Multi-Objective Optimization (BRIDGE-MO), an optimizer that represents alternative routes to joint attainment through a compact frontier of witness points. An LLM proposes structured intervention cards conditioned on these witnesses, while posterior bottleneck screening and constrained reweighting guide candidate selection with independent exploration and real-evaluation feedback. We derive a conditional, one-step Gaussian-process certificate for risk-gated candidates under an idealized posterior model. Across six finite-pool scientific optimization tasks spanning two to eleven objectives, BRIDGE-MO achieves strong joint-margin trajectories and competitive target-attainment and Pareto performance, supported by module ablations and model-sensitivity analyses. Our results suggest that BRIDGE-MO provides a practical and theoretically grounded framework for scientific optimization focused on simultaneous target satisfaction instead of trade-off expansion.

1 INTRODUCTION

Many scientific design problems involve multiple properties that must satisfy explicit performance requirements simultaneously. For example, molecular design may jointly constrain electronic, thermochemical, and structural properties, while photovoltaic design may impose requirements on several performance metrics (Ramakrishnan et al., 2014; Nagasawa et al., 2018). In such settings, the design objective is to identify candidates that simultaneously meet all prescribed property targets, rather than merely achieve favorable trade-offs among competing objectives. We refer to this setting as *fixed-target multi-objective joint-attainment optimization*. The challenge is particularly acute when each evaluation requires an expensive experiment or high-fidelity simulation and only a limited evaluation budget is available.

Multi-objective optimization seeks effective trade-offs among competing objectives. In expensive black-box settings, multi-objective Bayesian optimization (MOBO) uses probabilistic surrogates and uncertainty-aware acquisition rules to guide evaluations (Shahriari et al., 2016; Knowles, 2006; Daulton et al., 2020; 2021a;b). Most classical methods follow the Pareto principle, seeking solutions for which improving one objective requires sacrificing another. This is well suited to characterizing alternative trade-offs, but a strong Pareto solution can still fail a fixed-target task because a single critical property remains below threshold.

This distinction creates a structural mismatch between Pareto-oriented search and target-oriented optimization. Pareto criteria reward relative improvement and frontier expansion, whereas joint attainment depends on absolute threshold satisfaction. Once an objective reaches its target, further improvement may contribute little to feasibility while an unresolved objective remains limiting. Target-oriented methods incorporate prescribed values or aspiration levels (Hanaoka, 2021; Hase

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

et al., 2018), but specifying targets alone does not determine how multiple near-feasible states should be preserved during sequential search. Different candidates may be limited by different objectives and therefore admit distinct routes to feasibility. **How can an optimizer preserve and exploit these complementary routes toward joint attainment under a limited evaluation budget?**

This formulation reveals a natural *bottleneck structure*: because joint attainment is determined by the minimum target margin, feasibility is governed by the least satisfied objective. Crucially, candidates with similar bottleneck values can exhibit different deficit profiles and therefore require different routes toward feasibility. Concentrating the search around a single incumbent can prematurely discard such complementary opportunities. We instead maintain a compact *multi-witness frontier*, in which each experimentally evaluated witness preserves a distinct near-feasible state and represents an alternative route toward joint attainment.

Building on this frontier, an LLM proposes structured interventions for individual witnesses, while a posterior lower-tail risk functional determines how far each route can be credibly advanced toward the joint target. Language guidance is therefore revisable and path-specific, whereas posterior uncertainty and real evaluations determine whether a proposed route remains credible.

Based on this formulation, we introduce Bayesian Risk-Aware Intervention-Driven Guided Exploration for Multi-Objective Optimization (BRIDGE-MO). BRIDGE-MO combines a multi-witness frontier, posterior lower-tail risk, and witness-conditioned intervention cards to guide alternative routes toward joint attainment. An independent exploration channel and real-evaluation feedback prevent the search from becoming dependent on any single intervention hypothesis.

We establish a one-step GP confidence certificate for risk-gated candidates and a conditional expected-certificate improvement for KL-constrained language guidance aligned with certified repair. We evaluate the framework on six finite-pool tasks spanning two to eleven objectives, with comparisons and ablations focused on joint-target attainment.

The main contributions of this paper are summarized as:

- We formulate fixed-target multi-objective optimization through a *multi-witness frontier* motivated by the bottleneck structure of joint attainment. The frontier preserves complementary near-feasible states as distinct routes toward simultaneous target satisfaction rather than collapsing progress into a single incumbent or Pareto ranking.
- We develop a risk-aware witness-path optimization mechanism in which posterior bottleneck lower-tail risk determines the credible continuation of each witness route, while structured LLM intervention cards provide witness-specific and revisable guidance.
- We provide conditional guarantees for Gaussian-process (GP) confidence-based path certificates and for expected certificate reduction under aligned, KL-constrained card reweighting, and evaluate the framework on six scientific candidate pools spanning two to eleven objectives.

2 Related Work

Multi-objective Bayesian optimization. Bayesian optimization (BO) optimizes expensive black-box functions using probabilistic surrogates and uncertainty-aware acquisition rules (Snoek et al., 2012; Shahriari et al., 2016). Multi-objective BO extends this framework to competing objectives through scalarization, Pareto-based acquisition, and local search strategies (Knowles, 2006; Daulton et al., 2020; 2021a;b). These methods are effective for recovering high-quality trade-offs and diverse Pareto fronts, which remain central evaluation targets in much of the MOBO literature.

Target-oriented optimization. Target-oriented methods optimize toward prescribed values, aspiration levels, or desired regions in objective space rather than unrestricted Pareto expansion. Goal-oriented inverse design and achievement scalarization are representative examples (Hanaoka, 2021; Hase et al., 2018). Such formulations are particularly relevant when scientific or engineering success is defined by explicit specifications. As the number of objectives grows, however, simultaneous satisfaction becomes increasingly difficult because overall feasibility can be dominated by only a few limiting properties.

LLMs and learned optimizers. Large language models and pretrained sequence models have recently been incorporated into optimization for candidate generation, evolutionary search, and learned decision making (Liu et al., 2023; Hung et al., 2025; Zhang et al., 2026; Schwanke et al., 2026; Chen et al., 2025; Yu et al., 2025; Abhyankar et al., 2025; Gao et al., 2026). These studies illustrate how prior knowledge can be incorporated into search under limited observations. At the same time, model-generated proposals and learned policies are distinct from calibrated experimental evidence, which remains important when optimization decisions must account for uncertainty and feedback from real evaluations.

3 METHOD

3.1 PRELIMINARIES

Fixed-target multi-objective optimization. Let $\mathcal{X}$ denote a finite candidate pool with M objectives $f_1, \ldots, f_M$, where evaluating a candidate $x \in \mathcal{X}$ reveals $\mathbf{y}(x) = \mathbf{f}(x) + \boldsymbol{\epsilon}(x)$ at a cost. For each objective j, we specify a fixed target τ_j, an optimization direction $d_j \in \{-1, +1\}$, and a positive scale s_j. We define the normalized target margin as $z_j(x) = \frac{\big(f_j(x) - \tau_j\big) d_j}{s_j}$ with $j = 1, \ldots, M$, so that $z_j(x) \geq 0$ indicates that objective j satisfies its target. The overall feasibility of a candidate is determined by its joint bottleneck $b(x) = \min_{j \in [M]} z_j(x)$ and joint attainment occurs when $b(x) \geq 0$. Thus, optimization is governed by the most limiting objective rather than by improvement on already-satisfied ones. Pareto quality and hypervolume remain useful secondary diagnostics, but the primary objective is simultaneous target satisfaction.

Joint probabilistic surrogate. Before evaluation round t, let $\mathcal{D}_{t-1}$ denote the set of completed real evaluations. We model each normalized objective with a Gaussian-process posterior (Rasmussen & Williams, 2006),

$$z_j \mid \mathcal{D}_{t-1} \sim \mathcal{GP}\big(\mu^z_{t-1,j}, k_{t-1,j}\big), \qquad z_j(x) \mid \mathcal{D}_{t-1} \sim \mathcal{N}\Big(\mu^z_{t-1,j}(x), \big[\sigma^z_{t-1,j}(x)\big]^2\Big). \tag{1}$$

we couple marginal GP samples using a shrunk residual-correlation estimate (Ledoit & Wolf, 2004), preserving marginal uncertainty without specifying a multi-output kernel. The resulting joint samples $\mathbf{z}_t^{(s)}(x)$ support bottleneck-risk assessment and intervention screening. Only real evaluations update the surrogate.

3.2 MULTI-WITNESS FRONTIER

Fixed-target joint attainment is governed by a bottleneck effect: a candidate is only as successful as its least satisfied objective. Consequently, two candidates with similar overall progress may be limited by different objectives and therefore require different routes toward feasibility. To preserve these alternative routes, we retain a compact set of high-quality evaluated candidates with distinct target-margin profiles as *witnesses*, each providing a reference state for subsequent search.

Definition 1 (Witness and multi-witness frontier)**.** *At round t, a* witness *is an experimentally evaluated candidate $w \in \mathcal{D}_t$ whose target-margin profile $\mathbf{z}_t^{\mathrm{obs}}(w)$ represents a useful near-feasible state for continuing the search toward joint attainment. A* multi-witness frontier *is a compact set $\mathcal{W}_t = \{w_{t,1}, \ldots, w_{t,K_t}\}$ with $K_t \leq K_{\max}$. The set preserves high-quality and sufficiently distinct witness states, so that different elements of $\mathcal{W}_t$ correspond to complementary routes toward the joint target.*

Witness quality is determined directly from observed target margins. Let $z^{\mathrm{obs}}_{t,j}(x)$ denote the normalized observed margin of objective j. We define the observed bottleneck quality as

$$Q_t^{\mathrm{obs}}(x) = \min_{j \in [M]} z^{\mathrm{obs}}_{t,j}(x), \qquad w_{t,1} = \arg\max_{x \in \mathcal{D}_t} Q_t^{\mathrm{obs}}(x). \tag{2}$$

The primary witness is therefore the evaluated point with the strongest worst-objective margin. Additional witnesses are selected from high-quality evaluated candidates while encouraging separation from already retained routes. This quality–diversity criterion preserves distinct deficit patterns rather than multiple variants of the same bottleneck configuration.

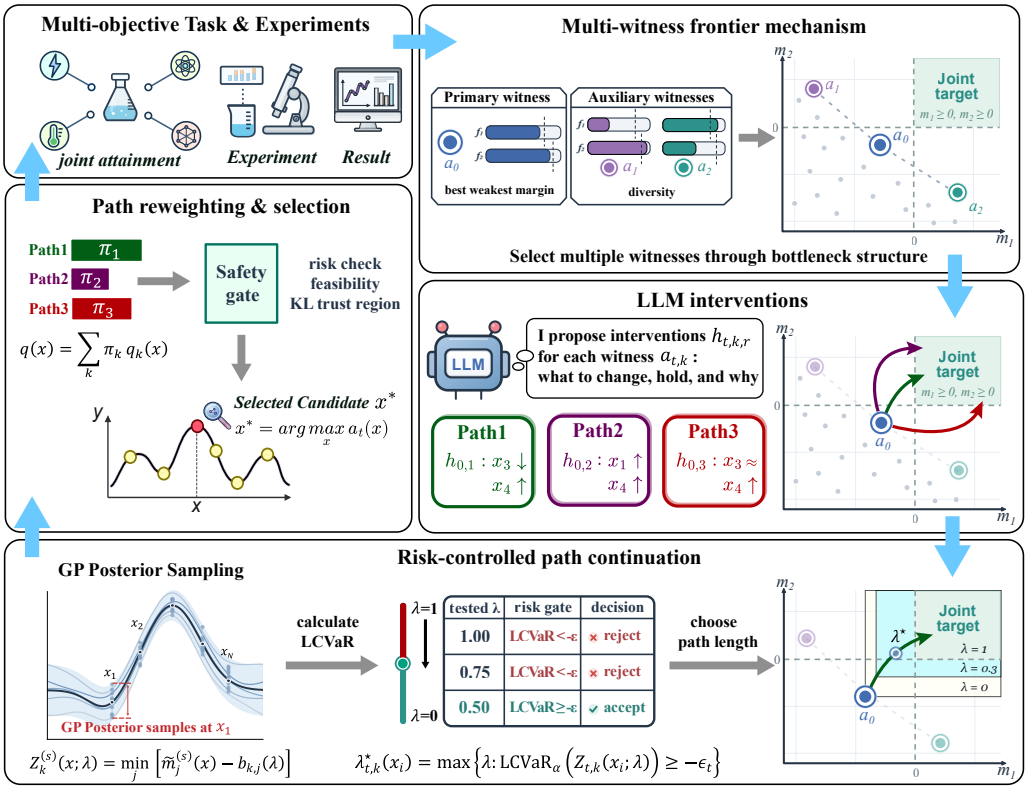

Figure 1: Overview of BRIDGE-MO. Real evaluations determine the posterior and multi-witness frontier. Witness-conditioned LLM cards propose local interventions, while lower-tail posterior risk determines credible progress toward joint attainment. Constrained reweighting and an independent exploration channel produce the final candidate distribution.

For subsequent path construction, each witness is associated with a conservative posterior margin

$$m^{L}_{t,k,j} = \mu^{z}_{t-1,j}(w_{t-1,k}) - \sqrt{\beta_t}\,\sigma^{z}_{t-1,j}(w_{t-1,k}). \tag{3}$$

which accounts for uncertainty around its current state. Witness membership is determined only by real evaluations, whereas the posterior margin is used to quantify how safely each experimentally supported route can be continued toward the joint target.

3.3 Witness-conditioned LLM intervention paths

Each witness provides a local scientific context for reasoning about how its remaining deficits may be repaired. Rather than asking an LLM to directly select the next experiment, BRIDGE-MO queries it for structured *intervention cards* associated with individual witness paths:

$$\mathcal{H}_{t,k} = \{h_{t,k,r}\}_{r=1}^{R_{t,k}}, \qquad h = (\mathcal{V}_h, \delta_h, \mathcal{A}_h, \mathcal{P}_h, c_h). \tag{4}$$

Here, $\mathcal{V}_h$ specifies the variables involved, δ_h their suggested directions or modes of change, $\mathcal{A}_h$ the objectives to improve, $\mathcal{P}_h$ the objectives to protect, and c_h the local applicability condition. A card therefore represents a directional and testable hypothesis for moving away from one witness toward joint attainment.

For candidate x, the adherence score $A_{t,k,r}(x) \in [0,1]$ measures consistency with the corresponding intervention. Cards are validated for schema, variable ranges, witness ownership, and candidate compatibility before use and remain path-specific. They affect only the decision layer, without creating pseudo-observations or modifying surrogate labels. After real evaluations, card reliability is updated based on adherence, improvement in limiting objectives, and preservation of protected objectives.

3.4 LOWER-TAIL RISK FUNCTIONAL

A candidate should not be favored merely because it follows an intervention card; its proposed progress must also remain credible under posterior uncertainty. For witness $w_{t,k}$, we construct a continuous family of intermediate target boundaries

$$b_{t,k,j}(\lambda) = (1-\lambda)\min\{m^L_{t,k,j}, 0\}, \qquad \lambda \in [0,1], \tag{5}$$

where $\lambda = 0$ corresponds to preserving the current witness state and $\lambda = 1$ corresponds to the final joint target. Objectives that are already satisfied remain protected at the target boundary.

Given a joint posterior sample $\mathbf{z}_t^{(s)}(x)$, we measure the weakest margin above this intermediate boundary by

$$Z^{(s)}_{t,k}(x;\lambda) = \min_{j\in[M]} \left[z^{(s)}_{t,j}(x) - b_{t,k,j}(\lambda) \right]. \tag{6}$$

The minimum preserves the bottleneck structure: strong performance on other objectives cannot compensate for a single severe failure.

We assess this margin using the lower-tail conditional value at risk (LCVaR), denoted by LCVaR_α:

$$\Phi_{t,k}(x;\lambda) = \mathrm{LCVaR}_\alpha \left(Z_{t,k}(x;\lambda) \mid \mathcal{D}_{t-1} \right), \tag{7}$$

using the lower-tail utility analogue of CVaR (Rockafellar & Uryasev, 2000),which summarizes the candidate's behavior in the least favorable posterior scenarios. Because Z is a margin, larger values are preferable. Candidates whose apparent progress relies on optimistic posterior samples therefore receive lower risk utility.

Finally, we define the maximum risk-feasible continuation level

$$\lambda^*_{t,k}(x) = \max\left\{ \lambda \in [0,1] : \Phi_{t,k}(x;\lambda) \geq -\epsilon_t \right\}. \tag{8}$$

Hence, $\lambda^*_{t,k}(x)$ directly measures how far candidate x can credibly advance witness path k toward the final joint target. Candidates with larger λ^* are preferred, while a bounded local-step or protection penalty can be used to distinguish candidates with similar continuation levels. The same lower-tail criterion also screens LLM interventions, so a card receives meaningful influence only when its proposed direction is supported by posterior evidence. Additional implementation details are provided in Appendix B.

3.5 OPTIMIZATION WORKFLOW

Figure 1 summarizes the BRIDGE-MO workflow. At each round, completed real evaluations are used to update the GP and refresh the multi-witness frontier. The LLM proposes witness-specific intervention cards, while posterior lower-tail risk determines each candidate's maximum risk-feasible continuation level along each path. Candidate selection combines these continuation scores with the adherence scores and reliability weights of validated cards.

Language-derived scores reweight the card-free distribution within a KL budget (Kullback & Leibler, 1951), $D_{\mathrm{KL}}(p_{\eta,t,k} \| p_{0,t,k}) \leq \kappa_{\mathrm{KL}}$. Each path distribution is mixed with an LLM-independent exploration distribution, and the resulting distributions are aggregated for diversity-aware batch selection. The selected batch is evaluated, and only real outcomes are appended to the dataset. These observations support subsequent GP and witness-frontier updates and revise card reliability, downweighting or refreshing unsupported interventions.

4 THEORETICAL GUARANTEES

We analyze a single decision at a fixed completed real-data history, holding the GP posterior and witness frontier fixed. The analysis connects GP-UCB confidence, risk-feasible continuation, and LLM-guided reweighting. Let $g(x) = \min\{b(x), 0\}$, $g^* = \max_{x\in\mathcal{X}} g(x)$, and $r_t(x) = g^* - g(x)$ denote the true one-step joint-utility regret.

Assume a simultaneous GP confidence event of probability at least $1-\delta$, on which $|z_j(x) - \mu^z_{t-1,j}(x)| \leq \sqrt{\beta_t}\sigma^z_{t-1,j}(x)$ for every candidate and objective. Let u_t be the clipped minimum of the coordinate-wise upper confidence bounds. The bottleneck GP-UCB reference selects

$x_t^0 \in \arg\max_{x\in\mathcal{X}} u_t(x)$, with confidence bound $B_t^0 = 2\sqrt{\beta_t}\max_j \sigma_{t-1,j}^z(x_t^0)$ from the GP-UCB argument (Srinivas et al., 2010).

For witness path k, let $d_{t,k} = \max_j[-m_{t,k,j}^L]_+$ be its conservative bottleneck deficit, where $[a]_+ = \max\{a, 0\}$. For exact lower-tail screening with Gaussian GP marginals, let $e_t(x)$ be the nonnegative risk-to-confidence slack defined in Appendix A.5. It accounts for the screening tolerance and the difference between posterior tail risk and GP confidence. For an accepted candidate, write $\lambda_t = \lambda_{t,k}^*(x_t)$ and $e_t = e_t(x_t)$. Its *net certified repair* is $V_{t,k}(x) = \lambda_{t,k}^*(x)d_{t,k} - e_t(x)$.

To isolate language guidance, fix a card-free path distribution p_0 with nonempty support consisting only of accepted candidates. The bounded aggregate card score U induces exponential reweighting $p_\eta(x) \propto p_0(x)e^{\eta U(x)}$, with guidance strength $\eta \geq 0$ and KL constraint $D_{\mathrm{KL}}(p_\eta\|p_0) \leq \kappa_{\mathrm{KL}}$, where $\kappa_{\mathrm{KL}} > 0$. The alignment $a_{t,k} = \mathrm{Cov}_{p_0}(U, V_{t,k})$ measures whether cards favor larger net certified repair. The posterior, support, path certificates, and card scores remain fixed in this comparison.

Theorem 1 (GP-UCB comparison with aligned language guidance)**.** *On the stated confidence event, $r_t(x_t^0) \leq B_t^0$. For any candidate x_t satisfying the exact path gate $\Phi_{t,k}(x_t; \lambda_t) \geq -\varepsilon_{\mathrm{risk},t}$,*

$$r_t(x_t) \leq (1-\lambda_t)d_{t,k} + e_t = d_{t,k} - V_{t,k}(x_t). \tag{9}$$

If $0 < d_{t,k} \leq B_t^0$, then

$$r_t(x_t) \leq B_t^{\mathrm{out}}, \qquad B_t^{\mathrm{out}} = B_t^0 + e_t. \tag{10}$$

If additionally $V_{t,k}(x_t) > 0$, then

$$r_t(x_t) \leq B_t^{\mathrm{in}}, \qquad B_t^{\mathrm{in}} = d_{t,k} - V_{t,k}(x_t) < B_t^0. \tag{11}$$

Moreover, let $C_{t,k}$ be the range-based curvature constant defined in Appendix A.6. If $a_{t,k} > 0$, every KL-accepted step $0 < \eta \leq a_{t,k}/(2C_{t,k})$ satisfies

$$\mathbb{E}_{p_\eta} r_t(X) \leq \mathbb{E}_{p_\eta}[d_{t,k} - V_{t,k}(X)] \leq \mathbb{E}_{p_0}[d_{t,k} - V_{t,k}(X)] - \frac{\eta a_{t,k}}{2}. \tag{12}$$

The expectations are over candidate selection at the fixed history. Thus positively aligned guidance strictly lowers the expected regret certificate relative to the card-free path distribution.

Remark (Reading the bounds). The condition $0 < d_{t,k} \leq B_t^0$ concerns witness quality, whereas $a_{t,k} > 0$ concerns guidance alignment. Under the witness condition, positive net repair gives a certificate below B_t^0; positive alignment and a controlled step reduce the expected certificate. Validation alone does not ensure alignment. The card-free path distribution p_0 is distinct from the deterministic GP-UCB reference x_t^0, and tighter certificates do not establish an ordering of their true regrets or cumulative performance. Appendix A.9 treats exploration, path aggregation, and the limitations of diversity postprocessing; Appendix A.8 provides the correction required for sampled-risk implementations.

5 Experiments

5.1 Experimental Settings

Baselines. We select representative methods from target-oriented optimization, Pareto-based multi-objective optimization, and LLM-assisted optimization as baselines.

Hanaoka (Hanaoka, 2021) directly targets the probability of reaching prescribed objective values and provides the closest target-oriented baseline.

qParEGO (Knowles, 2006) performs multi-objective optimization through randomized scalarization.

qLogNEHVI (Daulton et al., 2021a; Ament et al., 2023) uses a numerically stabilized noisy hypervolume-improvement acquisition function for Pareto-front optimization.

MORBO (Daulton et al., 2021b) performs multi-objective Bayesian optimization through coordinated local trust regions.

NSGA-II (Deb et al., 2002) is a classical evolutionary multi-objective optimizer based on nondominated sorting and crowding distance.

MOEA/D-LLM (Liu et al., 2023) incorporates an LLM as a search operator within decomposition-based multi-objective evolution.

BOFormer (Hung et al., 2025) uses a pretrained sequence model to learn non-Markovian acquisition behavior for Bayesian optimization.

TAMO (Zhang et al., 2026) amortizes multi-objective optimization with a pretrained transformer policy.

Implementation Details. Under each task's fixed-pool protocol, all methods use the same five initial candidates and a budget of 100 candidate evaluations; after initialization, candidates are selected in batches of $q = 4$. Each available method–task pair is summarized over the completed runs in its result package. Objective directions and thresholds are fixed before search; normalization is used only to fit the surrogate and to compute normalized joint margins.

The primary metric is the best-so-far joint margin, $\max_{i \le n} \min_j z_j(x_i)$. We also report first joint hit and success rate where applicable, with hypervolume as a secondary measure of Pareto quality. For clarity, the primary benchmark and the ablation studies use separate QM9 objective protocols; the complete protocol specification, including the shared initialization and the exact role of LLM warm-start, is provided in Appendix D.1.

Tasks. We evaluate all methods on six finite-pool scientific optimization benchmarks covering materials discovery, molecular design, photovoltaic optimization, structured engineering design, and reaction optimization, with objective dimensionality ranging from two to eleven.

NMD18 Oxide Discovery (2 objectives). The NOMAD 2018 oxide pool (Sutton et al., 2019; Langer et al., 2022) optimizes formation energy and target-gap error over candidate oxide materials.

PFD Polymer Photovoltaics (4 objectives). The PFD benchmark is derived from experimentally collected polymer-photovoltaic measurements (Nagasawa et al., 2018) and jointly targets power-conversion efficiency, open-circuit voltage, short-circuit current density, and fill factor.

QM9 Molecular Discovery (11 objectives). QM9 (Ramakrishnan et al., 2014) provides a finite molecular candidate pool with eleven electronic, thermochemical, and molecular-property targets that must be satisfied jointly.

Sandwich (6 objectives). Sandwich is a deterministic 48-variable, six-objective design oracle inspired by the adequacy and moderation structure of HEI-2020 (Shams-White et al., 2023).

Tartarus GDB13 Emitter Discovery (2 objectives). This task is drawn from the Tartarus benchmark suite (Nigam et al., 2023) and optimizes two target properties over a molecular emitter pool.

Tartarus Reaction-Substrate Discovery (5 objectives). The second Tartarus task (Nigam et al., 2023) considers a finite reaction–substrate candidate pool with five prescribed targets.

5.2 Main results

BRIDGE-MO achieves the lowest mean first-hit count on all six tasks in Table 1. These results demonstrate consistent advantages in reaching simultaneous target satisfaction under the fixed evaluation budget. Detailed search trajectories are shown in Fig. 2.

Trajectory behavior. Figure 2 compares the evolution of the mean best-so-far joint margin across the six fixed-target tasks. BRIDGE-MO improves rapidly during the early search and reaches the joint-feasible region earlier on the tasks where the target constraints are most tightly coupled. On PFD and Sandwich, its trajectory remains above those of most baselines through a substantial portion of the evaluation budget, although several methods reduce the gap during later exploration. On both Tartarus tasks, BRIDGE-MO crosses the joint-attainment threshold earlier and continues to improve, whereas several competing methods remain constrained by a negative bottleneck margin at the end of the budget. These results indicate that the advantage of BRIDGE-MO is primarily an improvement in search direction and early target attainment, rather than merely a higher final margin. The trajectory-level pattern is consistent with the method's objective: it focuses the search on the currently limiting objective while preserving multiple candidate routes toward simultaneous feasibility.

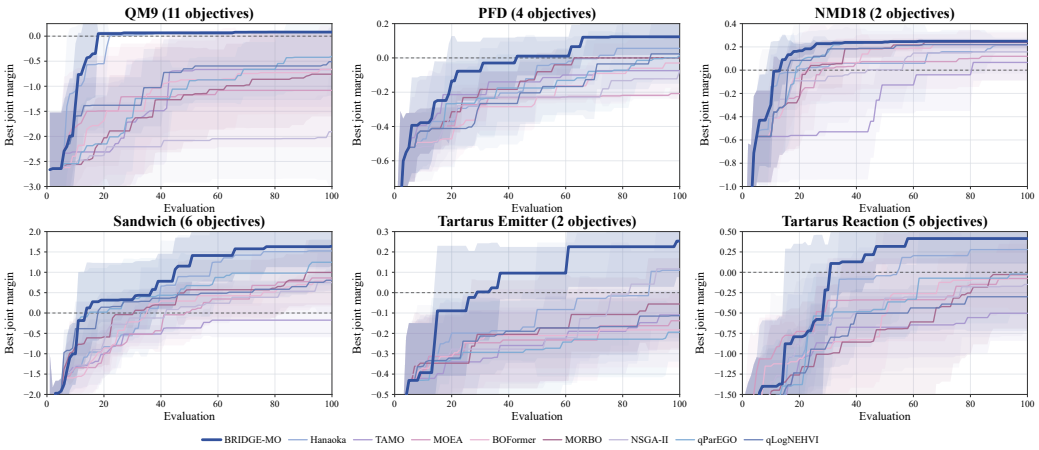

Figure 2: Main fixed-target comparison across six finite scientific pools. Panels are ordered as QM9 (11 objectives), PFD (4), NMD18 (2), Sandwich (6), Tartarus emitter (2), and Tartarus reaction (5). Curves show the mean best-so-far joint margin; shaded regions denote ± 1 standard deviation. Only methods with available traces are plotted for a given task.

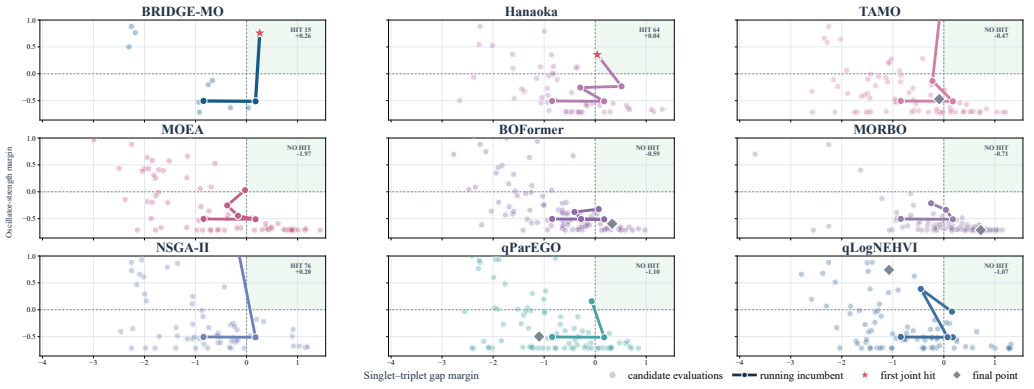

Figure 3: Search trajectories on Tartarus Emitter in two-dimensional target-margin space. Points denote evaluated candidates; the pale-green quadrant indicates joint feasibility. Dark lines connect successive bottleneck-improving incumbents, red stars mark first joint hits, and grey diamonds mark endpoints without joint attainment.

Search behavior. Figure 3 illustrates the search behavior on the Tartarus Emitter dataset, where the upper-right quadrant represents simultaneous satisfaction of both target constraints. Compared with the baselines, BRIDGE-MO exhibits a more directed incumbent progression toward this joint-feasible region, whereas other methods generate more dispersed candidate clouds and often remain limited by one of the two target margins. This behavior reflects the multi-witness mechanism: complementary witness paths preserve alternative routes toward joint attainment, while risk-aware aggregation focuses evaluations on candidates that improve the currently limiting target. The corresponding aggregate first-hit statistics across repeated experiments are reported in the main benchmark results.

5.3 Ablation and Model Sensitivity

Figure 4 gives two complementary views of the separate QM9 five-objective ablation: time to the first joint hit and final hypervolume. These curves are not the eleven-objective main benchmark.

Module ablation. Figure 4 shows complementary contributions from the ablated components. Removing batch diversity causes the largest first-hit delay, supporting the value of evaluating com-

Table 1: First joint-attainment evaluation under the fixed-target protocol. Values are mean ± standard deviation across the available seed traces, followed by the number of successful seeds $(x/5)$. An evaluated seed with no joint hit within the 100-candidate budget is coded as 101; unavailable traces are reported as missing rather than as failures.

Method	QM9 (11 objectives)	PFD (4 objectives)	NMD18 (2 objectives)	Sandwich (6 objectives)	Tartarus emitter (2 objectives)	Tartarus reaction (5 objectives)
BRIDGE-MO	$\mathbf{15.4 \pm 3.4}$ **(5/5)**	$\mathbf{41.6 \pm 24.5}$ **(5/5)**	$\mathbf{12.4 \pm 3.6}$ **(5/5)**	$\mathbf{21.4 \pm 11.8}$ **(5/5)**	$\mathbf{31.2 \pm 19.3}$ **(5/5)**	$\mathbf{33.4 \pm 12.5}$ **(5/5)**
Hanaoka	21.6 ± 10.9 (5/5)	67.8 ± 36.3 (3/5)	27.4 ± 34.3 (5/5)	42.6 ± 13.3 (5/5)	75.0 ± 21.9 (4/5)	54.2 ± 34.3 (4/5)
TAMO	70.6 ± 43.7 (2/5)	101.0 ± 0.0 (0/5)	67.8 ± 23.0 (4/5)	91.8 ± 20.6 (1/5)	101.0 ± 0.0 (0/5)	101.0 ± 0.0 (0/5)
MOEA	82.0 ± 42.5 (1/5)	101.0 ± 0.0 (0/5)	39.2 ± 36.3 (4/5)	60.0 ± 22.0 (5/5)	101.0 ± 0.0 (0/5)	69.6 ± 38.3 (3/5)
BOFormer	82.0 ± 39.7 (2/5)	92.2 ± 19.7 (1/5)	23.4 ± 13.2 (5/5)	28.8 ± 6.8 (5/5)	82.8 ± 40.7 (1/5)	84.2 ± 24.1 (2/5)
MORBO	89.0 ± 16.4 (2/5)	48.6 ± 12.6 (5/5)	24.8 ± 10.9 (5/5)	24.2 ± 13.4 (5/5)	86.2 ± 33.1 (1/5)	76.6 ± 28.9 (3/5)
NSGA-II	101.0 ± 0.0 (0/5)	90.2 ± 14.4 (3/5)	37.8 ± 41.0 (4/5)	55.2 ± 30.0 (4/5)	80.8 ± 19.2 (3/5)	72.6 ± 41.4 (2/5)
qParEGO	97.4 ± 7.5 (2/5)	67.6 ± 45.3 (3/5)	18.6 ± 6.3 (5/5)	30.4 ± 17.9 (5/5)	101.0 ± 0.0 (0/5)	67.6 ± 33.6 (3/5)
qLogNEHVI	87.6 ± 28.9 (2/5)	78.6 ± 29.8 (3/5)	18.0 ± 7.1 (5/5)	25.6 ± 17.0 (5/5)	101.0 ± 0.0 (0/5)	96.6 ± 9.8 (1/5)

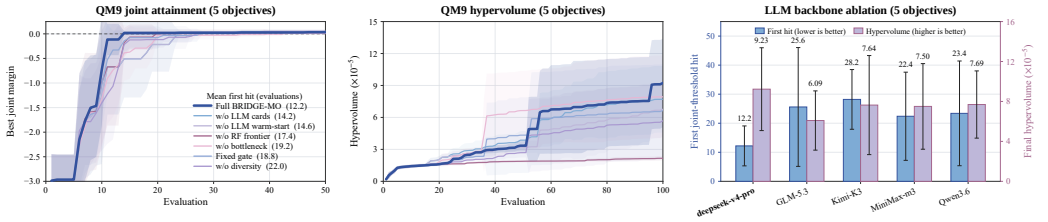

Figure 4: Main ablation on the separate QM9 five-objective protocol. The left and middle panels show first joint-attainment and final hypervolume trajectories for the retained module variants; the right panel reports the two endpoint metrics for the five retained language-model backbones.

plementary candidates. Removing bottleneck scoring or replacing adaptive gating with a fixed gate also delays attainment, consistent with their roles in identifying limiting objectives and regulating path expansion. Removing frontier risk filtering (*w/o RF frontier*) most strongly reduces final hypervolume and slows attainment, highlighting the importance of screening repair routes. Removing LLM cards or warm-start produces smaller delays, suggesting that language guidance complements the GP-driven search. Together, these results support the joint contribution of guided proposals, risk screening, and complementary candidate selection.

Model ablation. With the optimization settings fixed, BRIDGE-MO achieves joint attainment across all tested language backbones, demonstrating consistent effectiveness beyond a single model. **deepseek-v4-pro** provides the best mean attainment speed and final hypervolume, while the other backbones also support successful optimization. These results suggest that language models provide useful, interchangeable guidance within the shared GP, witness, and risk-screening framework. The overall joint-attainment capability is therefore preserved across backbone choices, supporting the robustness and flexibility of BRIDGE-MO.

6 Conclusion

BRIDGE-MO integrates witness-specific intervention cards with GP-based posterior risk screening and KL-constrained candidate reweighting for fixed joint-target optimization. Its multi-witness design preserves complementary repair routes, while bottleneck scoring, adaptive gating, and batch diversity focus evaluations on limiting objectives without sacrificing exploration coverage. Across six scientific candidate pools, BRIDGE-MO consistently accelerates joint target attainment and maintains strong Pareto quality. The theoretical results further establish GP-based certificates and conditional gains from aligned risk-aware reweighting, providing a principled foundation for efficient and reliable multi-objective search.

AI USE STATEMENT

Generative AI tools were used in two roles in this work. First, large language models were used as an experimental component of BRIDGE-MO to generate structured, witness-conditioned intervention cards from task descriptions and completed observations. These model outputs were used only as decision-layer guidance and were not treated as ground-truth measurements or used to create pseudo-observations for surrogate-model fitting. Second, generative AI tools were used to assist with manuscript drafting and language and technical editing. All AI-assisted content and model-generated outputs used in the study were reviewed by the authors. The authors take full responsibility for the final content of the paper, including all claims, analyses, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We provide detailed information to support reproducibility of both the theoretical and empirical results. Complete assumptions and proofs for the theoretical guarantees are included in the appendix. The implementation details of the multi-witness frontier, posterior risk screening, intervention card scoring, and real-evaluation feedback are also specified. For the experiments, we report the provenance and objective definitions of all benchmark tasks, the evaluated baselines, evaluation budgets, matched-seed protocols, ablation settings, and language-model sensitivity analyses. We further provide the structured prompt contract and validation rules used for LLM-generated intervention cards. These details are intended to enable independent reproduction of the reported optimization procedure and experimental comparisons.

REFERENCES

Nikhil Abhyankar, Sanchit Kabra, Saaketh Desai, and Chandan K. Reddy. LLEMA: Evolutionary Search with LLMs for Multi-Objective Materials Discovery. *arXiv preprint arXiv:2510.22503*, 2025. doi: 10.48550/arXiv.2510.22503. URL https://arxiv.org/abs/2510.22503.

Sebastian Ament, Samuel Daulton, David Eriksson, Maximilian Balandat, Eytan Bakshy, and Jasper Snoek. Unexpected improvements to expected improvement algorithms for bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Zhen-Song Chen, Hong-Wei Ding, Xian-Jia Wang, and Witold Pedrycz. LLM4CMO: Large Language Model-aided Algorithm Design for Constrained Multiobjective Optimization. *arXiv preprint arXiv:2508.11871*, 2025. doi: 10.48550/arXiv.2508.11871. URL https://arxiv.org/abs/2508.11871.

Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 844–853. PMLR, 2017.

Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. Differentiable expected hypervolume improvement for parallel multi-objective bayesian optimization. *Advances in Neural Information Processing Systems*, 2020.

Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. Parallel bayesian optimization of multiple noisy objectives with expected hypervolume improvement. In *Advances in Neural Information Processing Systems*, 2021a.

Samuel Daulton, David Eriksson, Maximilian Balandat, and Eytan Bakshy. Multi-objective bayesian optimization over high-dimensional search spaces. In *Advances in Neural Information Processing Systems*, volume 34, 2021b.

Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

Rui Gao, Youngseung Jeon, Swastik Roy, Morteza Ziyadi, and Xiang 'Anthony' Chen. C-MORAL: Controllable Multi-Objective Molecular Optimization with Reinforcement Alignment for LLMs. *arXiv preprint arXiv:2604.23061*, 2026. doi: 10.48550/arXiv.2604.23061. URL https://arxiv.org/abs/2604.23061.

Kyohei Hanaoka. Bayesian optimization for goal-oriented multi-objective inverse material design. *iScience*, 24(7):102781, 2021. doi: 10.1016/j.isci.2021.102781.

Florian Hase, Luis M. Roch, Christoph Kreisbeck, and Alán Aspuru-Guzik. Chimera: Enabling hierarchy based multi-objective optimization for self-driving laboratories. *Chemical Science*, 9: 7642–7655, 2018. doi: 10.1039/C8SC02239A.

Yu-Heng Hung, Kai-Jie Lin, Yu-Heng Lin, Chien-Yi Wang, Cheng Sun, and Ping-Chun Hsieh. BOFormer: Learning to solve multi-objective bayesian optimization via non-markovian RL. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=UnCKU8pZVe.

Joshua Knowles. Parego: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, 10:50–66, 2006.

Solomon Kullback and Richard A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951. doi: 10.1214/aoms/1177729694.

Matthias F. Langer, Alexander Goeßmann, and Matthias Rupp. Representations of molecules and materials for interpolation of quantum-mechanical simulations via machine learning. *npj Computational Materials*, 8:41, 2022. doi: 10.1038/s41524-022-00721-x.

Olivier Ledoit and Michael Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, 2004. doi: 10.1016/S0047-259X(03)00096-4.

Fei Liu, Xi Lin, Zhenkun Wang, Shunyu Yao, Xialiang Tong, Mingxuan Yuan, and Qingfu Zhang. Large language model for multi-objective evolutionary optimization. In *arXiv preprint arXiv:2310.12541*, 2023.

Pascal Massart. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The annals of Probability*, pp. 1269–1283, 1990.

Shinji Nagasawa, Eman Al-Naamani, and Akinori Saeki. Computer-aided screening of conjugated polymers for organic solar cell: Classification by random forest. *The Journal of Physical Chemistry Letters*, 9(10):2639–2646, 2018. doi: 10.1021/acs.jpclett.8b00635.

AkshatKumar Nigam, Robert Pollice, Gary Tom, Kjell Jorner, John Willes, Luca A. Thiede, Anshul Kundaje, and Alán Aspuru-Guzik. Tartarus: A benchmarking platform for realistic and practical inverse molecular design. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-0145.

Raghunathan Ramakrishnan, Pavlo O. Dral, Matthias Rupp, and O. Anatole von Lilienfeld. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific Data*, 1:140022, 2014.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

R. Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *The Journal of Risk*, 2(3):21–41, 2000. doi: 10.21314/JOR.2000.038.

R. Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7):1443–1471, 2002. doi: 10.1016/S0378-4266(02)00271-6.

Andrej Schwanke, Lyubomir Ivanov, David Salinas, Frank Hutter, and Arber Zela. Multi-Objective Hierarchical Optimization with Large Language Models. *arXiv preprint arXiv:2601.13892*, 2026. doi: 10.48550/arXiv.2601.13892. URL https://arxiv.org/abs/2601.13892.

Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1): 148–175, 2016.

Marissa M. Shams-White, T. E. Pannucci, J. L. Lerman, et al. Healthy eating index-2020: Review and update process to reflect the dietary guidelines for americans, 2020–2025. *Journal of the Academy of Nutrition and Dietetics*, 123:1280–1288, 2023.

Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, 2012.

Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on Machine Learning*, pp. 1015–1022, 2010.

Christopher Sutton, Luca M. Ghiringhelli, Takenori Yamamoto, Yury Lysogorskiy, Lars Blumenthal, Thomas Hammerschmidt, Jacek R. Golebiowski, Xiangyue Liu, Angelo Ziletti, and Matthias Scheffler. Crowd-sourcing materials-science challenges with the NOMAD 2018 Kaggle competition. *npj Computational Materials*, 5:111, 2019. doi: 10.1038/s41524-019-0239-3.

Jiajun Yu, Yizhen Zheng, Huan Yee Koh, Shirui Pan, Tianyue Wang, and Haishuai Wang. Collaborative Expert LLMs Guided Multi-Objective Molecular Optimization. *arXiv preprint arXiv:2503.03503*, 2025. doi: 10.48550/arXiv.2503.03503. URL https://arxiv.org/abs/2503.03503.

Xinyu Zhang, Conor Hassan, Julien Martinelli, Daolang Huang, and Samuel Kaski. In-context multi-objective optimization. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/901b378426e1816eabb776a786c02f41-Abstract-Conference.html.

A Full Theoretical Details

This appendix proves Theorem 1. The argument retains four certificate steps: simultaneous marginal GP confidence, a bottleneck GP-UCB reference bound, conversion of the posterior path gate into true margin bounds, and comparison of the resulting one-step certificates. It additionally connects LLM card reweighting to the expected certificate through alignment with net certified repair. All comparisons use the same completed history. We neither modify the GP prior through a lift nor assume that the minimum of several GP objectives is itself a GP. The final subsections treat numerical screening and distinguish the conditions needed for pointwise certificates from those needed for an expected-certificate improvement after final selection.

A.1 Setting, probability conventions, and marginal confidence

Let $\mathcal{X}$ be a nonempty finite pool, $N = |\mathcal{X}|$, and let $M \geq 1$. The true normalized margins $z_j(x)$ use the fixed targets, directions, and positive scales in Section 3. Set

$$b(x) = \min_{j\in[M]} z_j(x), \qquad g(x) = \min\{b(x), 0\}, \qquad g^* = \max_{x\in\mathcal{X}} g(x), \qquad r_t(x) = g^* - g(x).$$

Thus $r_t(x) \geq 0$ and $g^* \leq 0$. If a feasible point exists, then $g^* = 0$ and $r_t(x) = -g(x)$ is exactly its joint-target deficit. Otherwise, regret is measured against the best available infeasible utility. These are latent quantities, not noisy observed margins.

Before round t, $\mathcal{D}_{t-1}$ contains only completed real evaluations and $\mathcal{W}_{t-1} \subseteq \mathcal{D}_{t-1}$ is the current witness set. The posterior mean and latent standard deviation for objective j are $\mu^z_{t-1,j}$ and $\sigma^z_{t-1,j}$. For a witness $w_{t-1,k}$,

$$m^L_{t,k,j} = \mu^z_{t-1,j}(w_{t-1,k}) - \sqrt{\beta_t}\sigma^z_{t-1,j}(w_{t-1,k}), \qquad d_{t,k} = \max_j[-m^L_{t,k,j}]_+. \tag{13}$$

The same pre-evaluation posterior and β_t are used throughout. In particular, no outcome of the candidate being certified is used to construct its certificate.

A sufficient statistical model. One sufficient Bayesian model takes independent latent GPs across objectives, fixed prior means and kernels, and independent observation errors $\epsilon^z_{n,j} \sim \mathcal{N}(0, \nu_j^2)$ with fixed $\nu_j^2 > 0$. Observations are $y^z_{n,j} = z_j(x_n) + \epsilon^z_{n,j}$; errors are independent of the latent functions. The targets, normalization, prior specification, and any task side information are fixed before response data are observed. Selection may adapt to completed evaluations and independent random seeds, but it does not use unobserved responses or additional unmodeled observations of the latent objectives. Within a synchronous batch, no selection uses that batch's outcomes. These conditions make the stated per-objective posterior exact under this sufficient model. Correlation introduced solely for risk scoring below is not treated as a change to this evidence model.

Lemma 1 (Simultaneous finite-pool confidence)**.** *Under the sufficient Gaussian model, let* $\delta \in (0, 1)$ *and*

$$\beta_t = 2 \log\left(\frac{MN\pi^2 t^2}{3\delta}\right). \tag{14}$$

Then the event

$$\mathcal{E}_\delta = \left\{ |z_j(x) - \mu^z_{t-1,j}(x)| \le \sqrt{\beta_t}\sigma^z_{t-1,j}(x) \text{ for every } t \ge 1,\ j \in [M],\ x \in \mathcal{X} \right\} \tag{15}$$

has probability at least $1 - \delta$.

Proof. Conditional on a completed adaptive history, a nondegenerate standardized posterior error is standard normal. Its two-sided tail at $\sqrt{\beta_t}$ is at most $2e^{-\beta_t/2}$. A zero posterior variance implies zero error almost surely under this model. Taking expectations and a union bound over t, j, x gives

$$\Pr(\mathcal{E}_\delta^c) \le \sum_{t=1}^{\infty} MN\, 2e^{-\beta_t/2} = \frac{6\delta}{\pi^2}\sum_{t=1}^{\infty} t^{-2} = \delta.$$

This is the finite-pool GP-UCB confidence argument (Srinivas et al., 2010). □

The GP and risk certificates below are deterministic implications of the confidence event at the analyzed history. Hence they also apply under any other model that justifies this event. The finite-distribution comparisons for card reweighting do not themselves require GP correctness. An RKHS model requires its own norm, noise, regularization, and confidence-radius assumptions (Chowdhury & Gopalan, 2017); the Gaussian schedule above cannot simply be reused. Data-dependent hyperparameter refits or an arbitrary numerical nugget do not establish confidence coverage. No RKHS assumption is made on the clipped bottleneck g.

The probability statement concerns the experiment before its history is realized. Fixing a realized history for the algebra does not assert that $\Pr(\mathcal{E}_\delta \mid \mathcal{D}_{t-1}) \ge 1 - \delta$ for every such history. For a single fixed history under the exact Gaussian model, the choice $\beta = 2\log(2MN/\delta)$ instead gives a directly conditional, all-candidate confidence statement by the same finite union bound.

A.2 Bottleneck envelopes and the GP-UCB reference

The coordinate-wise lower bound is $L_{t,j}(x) = \mu^z_{t-1,j}(x) - \sqrt{\beta_t}\sigma^z_{t-1,j}(x)$. The corresponding upper bound is $U_{t,j}(x) = \mu^z_{t-1,j}(x) + \sqrt{\beta_t}\sigma^z_{t-1,j}(x)$. Their clipped bottlenecks are

$$\ell_t(x) = \min\{0, \min_j L_{t,j}(x)\}, \qquad u_t(x) = \min\{0, \min_j U_{t,j}(x)\}. \tag{16}$$

Lemma 2 (Bottleneck confidence envelopes)**.** *On* $\mathcal{E}_\delta$*, for every candidate,*

$$\ell_t(x) \le g(x) \le u_t(x), \qquad 0 \le u_t(x) - \ell_t(x) \le 2\sqrt{\beta_t}\max_j \sigma^z_{t-1,j}(x). \tag{17}$$

Proof. For any $a, c \in \mathbb{R}^M$, the inequalities $c_j - \|a - c\|_\infty \le a_j \le c_j + \|a - c\|_\infty$ imply $|\min_j a_j - \min_j c_j| \le \|a - c\|_\infty$. The map $v \mapsto \min\{v, 0\}$ is monotone and 1-Lipschitz. Apply monotonicity to the coordinate confidence intervals and nonexpansiveness to their endpoints, whose coordinate differences are $2\sqrt{\beta_t}\sigma^z_{t-1,j}(x)$. □

Proposition 1 (One-step bottleneck GP-UCB bound)**.** *Fix a tie-breaking rule and choose* $x_t^0 \in \arg\max_{x\in\mathcal{X}} u_t(x)$*. On* $\mathcal{E}_\delta$*,*

$$r_t(x_t^0) \le B_t^0, \qquad B_t^0 = 2\sqrt{\beta_t}\max_j \sigma^z_{t-1,j}(x_t^0). \tag{18}$$

Proof. Let $x^* \in \arg\max_{x\in\mathcal{X}} g(x)$, which exists by finiteness. On the confidence event, $g^* = g(x^*) \le u_t(x^*) \le u_t(x_t^0)$. Thus

$$r_t(x_t^0) \le u_t(x_t^0) - g(x_t^0) \le u_t(x_t^0) - \ell_t(x_t^0) \le B_t^0.$$

This extends the single-step optimism argument of Srinivas et al. (2010, Lemma 5.2) to the clipped coordinate minimum. □

The reference is the specified bottleneck-UCB rule under the same posterior, not every BO algorithm. B_t^0 is a confidence bound, not the reference's actual regret or its tightest possible certificate. The one-step proof does not need an information-gain summation or a batch variance constant.

A.3 Working posterior and lower-tail risk

Let $\mathbf{Z}_t(x)$ denote a draw from the working joint distribution used for risk scoring at candidate x. Its marginals satisfy

$$Z_{t,j}(x) \sim \mathcal{N}\big(\mu^z_{t-1,j}(x), [\sigma^z_{t-1,j}(x)]^2\big)\,. \tag{19}$$

These are latent-objective marginals, not predictive observations with additional measurement noise. Their dependence may be arbitrary. For example, a positive semidefinite correlation matrix R_t with unit diagonal gives the marginal-preserving construction $\mathbf{Z}_t(x) = \boldsymbol{\mu}_t(x) + \mathrm{diag}(\boldsymbol{\sigma}_t(x))R_t^{1/2}\xi$, where $\xi \sim \mathcal{N}(0, I_M)$ and the vectors contain the corresponding pre-evaluation means and standard deviations. Its marginal variances are unchanged because $(R_t)_{jj} = 1$. This is a sufficient construction, not a requirement for a fully specified multi-output kernel.

For any integrable scalar Y, define

$$Q_v(Y) = \inf\{y : \Pr(Y \le y) \ge v\}, \qquad \mathrm{LCVaR}_\alpha(Y) = \frac{1}{\alpha}\int_0^\alpha Q_v(Y)\,dv, \quad 0<\alpha<1. \tag{20}$$

This is the lower-tail utility form of CVaR; it includes distributions with atoms (Rockafellar & Uryasev, 2002).

Lemma 3 (Lower-tail monotonicity, stability, and Gaussian value)**.** *If* $Y \le Y'$ *almost surely under a common coupling, then* $\mathrm{LCVaR}_\alpha(Y) \le \mathrm{LCVaR}_\alpha(Y')$*. If* $|Y - Y'| \le h$ *almost surely, their lower-tail risks differ by at most* h*. For* $Y \sim \mathcal{N}(\mu, \sigma^2)$*, including* $\sigma = 0$*,*

$$\mathrm{LCVaR}_\alpha(Y) = \mu - c_\alpha\sigma, \qquad c_\alpha = \frac{\varphi(\Phi_{\mathrm{N}}^{-1}(\alpha))}{\alpha}. \tag{21}$$

Proof. Almost-sure ordering implies the same ordering of quantiles, so integration proves monotonicity. If $Y' - h \le Y \le Y' + h$, quantile translation and monotonicity give the stability inequality. For $\sigma > 0$, substitute $v = \Phi_{\mathrm{N}}(s)$ to obtain

$$\frac{1}{\alpha}\int_0^\alpha [\mu + \sigma\Phi_{\mathrm{N}}^{-1}(v)]\,dv = \mu + \frac{\sigma}{\alpha}\int_{-\infty}^{\Phi_{\mathrm{N}}^{-1}(\alpha)} s\varphi(s)\,ds = \mu - \frac{\sigma}{\alpha}\varphi(\Phi_{\mathrm{N}}^{-1}(\alpha)).$$

The final equality uses $\varphi'(s) = -s\varphi(s)$. For $\sigma = 0$, the distribution is constant and the same formula holds. This is the lower-tail counterpart of the standard Gaussian CVaR identity (Rockafellar & Uryasev, 2000; 2002). □

This lemma does not identify the working dependence model with the true joint law. The certificate below needs exact Gaussian working marginals and valid true marginal confidence, not a calibrated working copula. The tail fraction α is not an additional failure probability in the resulting $1-\delta$ confidence statement.

A.4 WELL-DEFINED RISK-FEASIBLE CONTINUATION

For path k, the boundary and scalar working-posterior margin are

$$\begin{aligned} b_{t,k,j}(\lambda) &= (1-\lambda)\min\{m^L_{t,k,j}, 0\}, \quad 0 \le \lambda \le 1, \\ Z_{t,k}(x;\lambda) &= \min_j [Z_{t,j}(x) - b_{t,k,j}(\lambda)], \qquad \Phi_{t,k}(x;\lambda) = \mathrm{LCVaR}_\alpha(Z_{t,k}(x;\lambda)). \end{aligned} \tag{22}$$

The scalar minimum is integrable because its absolute value is bounded by the sum of the absolute coordinate margins.

Lemma 4 (Continuation existence and boundary contraction)**.** *For fixed t, k, x, the function $\lambda \mapsto \Phi_{t,k}(x;\lambda)$ is nonincreasing and $d_{t,k}$-Lipschitz. Moreover,*

$$\min_j b_{t,k,j}(\lambda) = -(1-\lambda)d_{t,k}. \tag{23}$$

For $\varepsilon_{\mathrm{risk},t} \ge 0$, the acceptable set $\{\lambda \in [0,1] : \Phi_{t,k}(x;\lambda) \ge -\varepsilon_{\mathrm{risk},t}\}$ is either empty or a closed initial interval. Its maximum exists whenever it is nonempty.

Proof. Each boundary is nondecreasing because $\min\{m^L_{t,k,j}, 0\} \le 0$. Hence every coordinate margin, its minimum, and its lower-tail risk are nonincreasing. Under the same draw $\mathbf{Z}_t(x)$,

$$|Z_{t,k}(x;\lambda) - Z_{t,k}(x;\lambda')| \le \max_j |b_{t,k,j}(\lambda) - b_{t,k,j}(\lambda')| = d_{t,k}|\lambda - \lambda'|.$$

Lemma 3 proves Lipschitz continuity of $\Phi_{t,k}$. Also, $\min_j \min\{m^L_{t,k,j}, 0\} = -d_{t,k}$, which gives (23). Continuity and monotonicity imply the stated interval property and attainment of its maximum. □

An empty acceptable set means rejection, denoted $\lambda^* = \perp$; it must not be identified with an acceptable level zero. It is empty exactly when the gate already fails at zero. The following results hold for any accepted level, including the exact maximum λ^* or an accepted grid level. They do not require an optimizer to find the maximum without numerical error. The contraction concerns conservative target boundaries; it is not an assertion that the observed witness state has physically evolved by that fraction.

A.5 FROM THE PATH GATE TO A TRUE REGRET CERTIFICATE

Proposition 2 (Risk-to-confidence conversion)**.** *Fix $x \in \mathcal{X}$, a witness path k, and $\lambda \in [0,1]$. Suppose the exact gate $\Phi_{t,k}(x;\lambda) \ge -\varepsilon_{\mathrm{risk},t}$ holds. Under* (19)*, on $\mathcal{E}_\delta$,*

$$z_j(x) \ge b_{t,k,j}(\lambda) - \varepsilon_{\mathrm{risk},t} - [\sqrt{\beta_t} - c_\alpha]_+ \sigma^z_{t-1,j}(x) \quad \textit{for every } j. \tag{24}$$

Consequently, writing

$$e_t(x) = \varepsilon_{\mathrm{risk},t} + [\sqrt{\beta_t} - c_\alpha]_+ \max_j \sigma^z_{t-1,j}(x), \tag{25}$$

we have

$$0 \le r_t(x) \le -g(x) \le (1-\lambda)d_{t,k} + e_t(x). \tag{26}$$

Proof. For each coordinate, $Z_{t,k}(x;\lambda) \le Z_{t,j}(x) - b_{t,k,j}(\lambda)$ under the same working joint distribution. Lemma 3 and the exact gate therefore imply

$$\begin{aligned} -\varepsilon_{\mathrm{risk},t} &\le \Phi_{t,k}(x;\lambda) \\ &\le \mathrm{LCVaR}_\alpha(Z_{t,j}(x) - b_{t,k,j}(\lambda)) \\ &= \mu^z_{t-1,j}(x) - b_{t,k,j}(\lambda) - c_\alpha \sigma^z_{t-1,j}(x). \end{aligned} \tag{27}$$

Rearrange and use the lower side of the true GP confidence interval:

$$\begin{aligned} z_j(x) &\ge \mu^z_{t-1,j}(x) - \sqrt{\beta_t}\sigma^z_{t-1,j}(x) \\ &\ge b_{t,k,j}(\lambda) - \varepsilon_{\mathrm{risk},t} - (\sqrt{\beta_t} - c_\alpha)\sigma^z_{t-1,j}(x) \\ &\ge b_{t,k,j}(\lambda) - \varepsilon_{\mathrm{risk},t} - [\sqrt{\beta_t} - c_\alpha]_+ \sigma^z_{t-1,j}(x). \end{aligned}$$

This proves (24). Replacing the coordinate-specific slack by (25) and taking minima gives $b(x) \ge -(1-\lambda)d_{t,k} - e_t(x)$. The right-hand side is nonpositive, so clipping $b(x)$ above at zero preserves the same lower bound for $g(x)$. Finally, $g^* \le 0$ gives $r_t(x) = g^* - g(x) \le -g(x)$, proving (26). □

Every inequality is pathwise on the same simultaneous confidence event. Consequently, the candidate, owning witness, and accepted level may be chosen adaptively from the current posterior and cards without an additional confidence union bound over these choices. Exact computation of the working risk is essential here; a finite empirical tail average requires the correction below.

A.6 Language guidance and net certified repair

A fixed-support comparison. Fix the completed history, witness path k, and all scores and certificates used in the comparison. Let p_0 be the card-free path distribution and $\mathcal{S} = \mathrm{supp}(p_0)$ its nonempty finite support. Every $x \in \mathcal{S}$ has an accepted level $\lambda^*_{t,k}(x) \in [0,1]$ under the exact path gate. For all candidates accepted on this path, including those in $\mathcal{S}$, define

$$V_{t,k}(x) = \lambda^*_{t,k}(x) d_{t,k} - e_t(x), \qquad \mathcal{B}_{t,k}(x) = d_{t,k} - V_{t,k}(x). \tag{28}$$

By Proposition 2, on $\mathcal{E}_\delta$,

$$0 \le r_t(x) \le \mathcal{B}_{t,k}(x) = (1 - \lambda^*_{t,k}(x)) d_{t,k} + e_t(x), \quad x \in \mathcal{S}. \tag{29}$$

The function $\mathcal{B}_{t,k}$ is introduced only to shorten the appendix proofs; it is exactly $d_{t,k} - V_{t,k}$ in the main text. Net certified repair $V_{t,k}$ is not an observed experimental improvement. It is the reduction of this certificate relative to the conservative starting deficit $d_{t,k}$, and it may be negative.

Let $U : \mathcal{S} \to \mathbb{R}$ be the actual finite aggregate card score. For $\eta \ge 0$, the exponential reweighting used in the main text is

$$p_\eta(x) = \frac{p_0(x) e^{\eta U(x)}}{\sum_{x' \in \mathcal{S}} p_0(x') e^{\eta U(x')}}, \qquad D_{\mathrm{KL}}(p_\eta \| p_0) \le \kappa_{\mathrm{KL}}. \tag{30}$$

The KL constraint is imposed on the accepted reweighting step. No hypothetical error-free LLM score is used. The comparison keeps $\mathcal{S}$, p_0, the posterior, U, and the accepted levels fixed while varying η. In particular, $V_{t,k}$ does not change with the reweighting strength. Since p_η and p_0 have the same support, no candidate rejected on this path is introduced. All expectations in this subsection are finite sums over these selection probabilities, with the realized history and true objective functions held fixed. They are not expectations over new measurements or posterior draws of the latent objective, and do not recondition the selection law on a future confidence event.

For $H : \mathcal{S} \to \mathbb{R}$, write $\mathrm{osc}_{\mathcal{S}}(H) = \max_{\mathcal{S}} H - \min_{\mathcal{S}} H$. Set

$$\begin{aligned} a_{t,k} &= \mathrm{Cov}_{p_0}(U, V_{t,k}), \\ D_U &= \mathrm{osc}_{\mathcal{S}}(U), \\ D_V &= \mathrm{osc}_{\mathcal{S}}(V_{t,k}), \\ C_{t,k} &= D_V D_U^2 / 8. \end{aligned} \tag{31}$$

Thus $C_{t,k}$ is an explicit finite-step remainder constant, not an unmeasured card-error term. Positive alignment implies $D_U, D_V > 0$ and therefore $C_{t,k} > 0$.

Lemma 5 (Directional change of the expected certificate)**.** *For every $\eta \ge 0$,*

$$\left| \mathbb{E}_{p_\eta} \mathcal{B}_{t,k} - \mathbb{E}_{p_0} \mathcal{B}_{t,k} + \eta a_{t,k} \right| \le C_{t,k} \eta^2. \tag{32}$$

In particular, if $a_{t,k} > 0$ and $0 < \eta \le a_{t,k}/(2C_{t,k})$, then

$$\mathbb{E}_{p_\eta} \mathcal{B}_{t,k} \le \mathbb{E}_{p_0} \mathcal{B}_{t,k} - \eta a_{t,k}/2. \tag{33}$$

Proof. Differentiating the finite normalizing sum gives

$$\frac{\mathrm{d}}{\mathrm{d}\eta} p_\eta(x) = p_\eta(x) \big(U(x) - \mathbb{E}_{p_\eta} U \big).$$

Consequently, for every fixed function H on $\mathcal{S}$,

$$\frac{\mathrm{d}}{\mathrm{d}\eta} \mathbb{E}_{p_\eta} H = \mathrm{Cov}_{p_\eta}(H, U).$$

This finite-support exponential-family identity follows directly by differentiating the normalizing constant. With $\Psi(\eta) = \mathbb{E}_{p_\eta} V_{t,k}$, it yields $\Psi'(0) = a_{t,k}$ and, on differentiating once more,

$$\Psi''(\eta) = \mathbb{E}_{p_\eta}\left[(V_{t,k} - \mathbb{E}_{p_\eta} V_{t,k})(U - \mathbb{E}_{p_\eta} U)^2\right]. \tag{34}$$

For verification, the expansion is $\mathbb{E}[V_{t,k}U^2] - 2\mathbb{E}[V_{t,k}U]\mathbb{E}[U] - \mathbb{E}[V_{t,k}]\mathbb{E}[U^2] + 2\mathbb{E}[V_{t,k}]\mathbb{E}[U]^2$, with all expectations under p_η.

For any random variable supported on an interval $[a, b]$, taking expectations of $(U - a)(b - U) \geq 0$ proves $\mathrm{Var}(U) \leq (\mathbb{E}U - a)(b - \mathbb{E}U) \leq (b - a)^2/4$. Hence

$$|\Psi''(\eta)| \leq D_V \,\mathrm{Var}_{p_\eta}(U) \leq D_V D_U^2/4 = 2C_{t,k}.$$

The integral Taylor remainder therefore gives

$$|\Psi(\eta) - \Psi(0) - \eta a_{t,k}| \leq \int_0^\eta (\eta - s) 2C_{t,k}\,\mathrm{d}s = C_{t,k}\eta^2.$$

Since $\mathbb{E}_{p_\eta}\mathcal{B}_{t,k} = d_{t,k} - \Psi(\eta)$, this proves (32). Under the stated step restriction, $C_{t,k}\eta^2 \leq \eta a_{t,k}/2$, so the upper side of that remainder bound proves (33). □

KL feasibility and effects of adverse alignment. Let $K(\eta) = \log \mathbb{E}_{p_0} e^{\eta U}$. Then $K'(\eta) = \mathbb{E}_{p_\eta} U$ and $K''(\eta) = \mathrm{Var}_{p_\eta}(U) \leq D_U^2/4$. Because $D_{\mathrm{KL}}(p_\eta \| p_0) = \eta K'(\eta) - K(\eta)$ vanishes at zero and has derivative $\eta K''(\eta)$, integration yields

$$D_{\mathrm{KL}}(p_\eta \| p_0) \leq \eta^2 D_U^2/8. \tag{35}$$

Thus for $a_{t,k} > 0$ and $\kappa_{\mathrm{KL}} > 0$ the explicit interval

$$0 < \eta \leq \min\left\{\frac{a_{t,k}}{2C_{t,k}}, \frac{\sqrt{8\kappa_{\mathrm{KL}}}}{D_U}\right\} \tag{36}$$

is nonempty and sufficient for both positive certificate improvement and KL feasibility. The main theorem permits any actually KL-accepted step meeting its first restriction; no largest finite accepted step is assumed to exist.

If $a_{t,k} < 0$, the lower side of (32) instead gives

$$\mathbb{E}_{p_\eta}\mathcal{B}_{t,k} \geq \mathbb{E}_{p_0}\mathcal{B}_{t,k} + \eta|a_{t,k}|/2, \quad 0 < \eta \leq |a_{t,k}|/(2C_{t,k}).$$

This is deterioration of the expected certificate, not a proof that the true expected regret deteriorates by the same amount. When $a_{t,k} = 0$, no direction is certified by the linear term. If $D_U = 0$, reweighting does not change the distribution; if $D_V = 0$, it cannot change the expected certificate. A singleton support has both degeneracies.

For any KL-accepted step, the following magnitude bound also holds:

$$|\mathbb{E}_{p_\eta}\mathcal{B}_{t,k} - \mathbb{E}_{p_0}\mathcal{B}_{t,k}| \leq D_V \min\{1, \sqrt{\kappa_{\mathrm{KL}}/2}\}. \tag{37}$$

Indeed, for finite distributions, the positive and negative parts of their difference each have mass equal to total variation. Hence an expectation difference is at most the function's oscillation times total variation. Pinsker's inequality gives $\|p_\eta - p_0\|_{\mathrm{TV}} \leq \sqrt{D_{\mathrm{KL}}(p_\eta\|p_0)/2}$. A self-contained justification is to apply the log-sum inequality on $A = \{x : p_\eta(x) > p_0(x)\}$ and its complement. The resulting binary relative entropy $d(a\|b)$ has second derivative $1/[a(1-a)] \geq 4$ in a and both value and first derivative zero at $a = b$; integrating twice gives $d(a\|b) \geq 2(a-b)^2$. Endpoint cases follow by limits. Combining these facts and $\mathrm{osc}_{\mathcal{S}}(\mathcal{B}_{t,k}) = D_V$ proves (37).

Interpretation of the alignment assumption. The covariance is computed from the actual score table and the fixed risk-certificate table. It measures whether the aggregate card score favors larger net certified repair. It is not a condition on witness quality $d_{t,k} \leq B_t^0$, a claim that validation establishes scientific correctness, or an assumption that the true regret has already improved. If posterior screening is itself part of the aggregate card score, positive covariance may partly reflect that shared model-based component; this theorem does not isolate a purely semantic contribution of the LLM. It neither assumes that every proposed card is correct nor proves that an arbitrary language model produces positive alignment.

A.7 Proof of the main theorem and comparison regimes

Proof of Theorem 1. Proposition 1 establishes the reference claim $r_t(x_t^0) \le B_t^0$. Apply Proposition 2 at $x = x_t$ and at the accepted λ_t. The value $e_t(x_t)$ is exactly e_t defined in the main text. This proves (9) without requiring a feasible point to exist.

If $0 < d_{t,k} \le B_t^0$, then $0 \le \lambda_t \le 1$ implies

$$(1-\lambda_t)d_{t,k} + e_t \le d_{t,k} + e_t \le B_t^0 + e_t = B_t^{\mathrm{out}},$$

which proves (10). If additionally $\lambda_t > e_t/d_{t,k}$, then

$$B_t^{\mathrm{in}} = (1-\lambda_t)d_{t,k} + e_t = d_{t,k} - (\lambda_t d_{t,k} - e_t) < d_{t,k} \le B_t^0.$$

Together with (9), this proves (11). The identity $V_{t,k}(x_t) = \lambda_t d_{t,k} - e_t$ is exactly the additional main-text notation for the same certificate.

For the LLM statement, apply Lemma 5 to the fixed actual card score and the net-repair table. Positive $a_{t,k}$ implies $C_{t,k} > 0$; the step condition gives $\mathbb{E}_{p_\eta}\mathcal{B}_{t,k} \le \mathbb{E}_{p_0}\mathcal{B}_{t,k} - \eta a_{t,k}/2$. On $\mathcal{E}_\delta$, the candidate-wise inequality $r_t(x) \le \mathcal{B}_{t,k}(x)$ holds for every point in the common support, so averaging proves the first inequality in (12). Substituting $\mathcal{B}_{t,k} = d_{t,k} - V_{t,k}$ proves its second inequality. All true-regret certificates use the same GP confidence event, while the score-expectation comparison is a deterministic identity and bound at the fixed history. The exact-risk theorem therefore needs no additional failure budget for the LLM step. □

What determines a tighter displayed bound. For $d_{t,k} > 0$, the exact algebraic condition for the path certificate to be smaller than B_t^0 is

$$(1-\lambda_t)d_{t,k} + e_t < B_t^0 \quad \Longleftrightarrow \quad \lambda_t > 1 - \frac{B_t^0 - e_t}{d_{t,k}}. \tag{38}$$

The main theorem gives the simpler sufficient condition $d_{t,k} \le B_t^0$ and $\lambda_t d_{t,k} > e_t$. In that regime the reduction in the displayed upper bound is

$$B_t^0 - B_t^{\mathrm{in}} = (B_t^0 - d_{t,k}) + \lambda_t d_{t,k} - e_t > 0.$$

The strict main-text condition is possible only if $e_t < d_{t,k}$. For $d_{t,k} = 0$, the general certificate becomes $r_t(x_t) \le e_t$; the ratio $e_t/d_{t,k}$ is not used. On $\mathcal{E}_\delta$, such a witness is already truly feasible. If $B_t^0 = 0$, the reference is certified optimal and no nonnegative regret certificate can be strictly smaller. If $d_{t,k} > B_t^0$, the general path certificate still holds, but the main-text B_t^{out} comparison is not available.

The slack separates the risk tolerance from the difference between two conservatism levels. If $c_\alpha \ge \sqrt{\beta_t}$, then $e_t = \varepsilon_{\mathrm{risk},t}$. Otherwise its second term is proportional to the largest candidate posterior standard deviation. A smaller tail fraction can increase c_α but also makes the exact gate more restrictive; it does not guarantee a larger feasible continuation level or a smaller combined certificate. The confidence multiplier cannot be reduced arbitrarily while retaining the same coverage claim.

Expected-certificate comparison with the GP-UCB width. For the fixed path distribution and an improving step,

$$\mathbb{E}_{p_\eta} r_t(X) \le d_{t,k} - \mathbb{E}_{p_0} V_{t,k} - \eta a_{t,k}/2.$$

Hence this displayed upper bound is below B_t^0 whenever

$$\mathbb{E}_{p_0} V_{t,k} + \eta a_{t,k}/2 > d_{t,k} - B_t^0. \tag{39}$$

A simpler sufficient condition is $d_{t,k} \le B_t^0$ and $\mathbb{E}_{p_0} V_{t,k} + \eta a_{t,k}/2 > 0$. Positive alignment alone does not imply either condition: a small improvement can leave the expected net repair negative. The reference p_0 here is a card-free path distribution, whereas x_t^0 is the deterministic bottleneck-UCB point. They are not identified. Since $\mathbb{E}_{p_0}\mathcal{B}_{t,k}$ is only an upper bound on $\mathbb{E}_{p_0} r_t$, decreasing the former does not imply a decrease of the latter. The improvement concerns the specific certificates.

True feasibility. If $\lambda_t = 1$ and $e_t = 0$, then (24) guarantees $z_j(x_t) \ge 0$ for every j. More generally, direct nonnegative coordinate lower bounds $L_{t,j}(x_t) \ge 0$ certify feasibility on $\mathcal{E}_\delta$. A small positive regret bound is not exact feasibility, and a favorable working-posterior probability is not a new experimental observation.

A.8 NUMERICAL SCREENING AND FINITE POSTERIOR SAMPLES

This subsection supplies an additional numerical-error extension, not an assumption that the current implementation already provides such a certificate. For S scalar posterior scenarios with ordered values $Y_{(1)} \le \cdots \le Y_{(S)}$, put $m = \lfloor \alpha S \rfloor$. The lower-tail risk of the empirical distribution is

$$\widehat{\mathrm{LCVaR}}_\alpha(Y) = \frac{\sum_{s=1}^{m} Y_{(s)} + (\alpha S - m) Y_{(m+1)}}{\alpha S}. \tag{40}$$

The empty sum is zero; a zero-weight final term is omitted. This formula includes the fractional sample weight and does not assume continuity.

Proposition 3 (Screening-error correction)**.** *Suppose an event $\mathcal{M}_\zeta$ of probability at least $1-\zeta$ provides, uniformly over all candidate–path–level triples eligible for selection,*

$$\widehat{\Phi}_{t,k}(x;\lambda) - \Phi_{t,k}(x;\lambda) \le \omega_{t,k}(x;\lambda), \qquad \omega_{t,k}(x;\lambda) \ge 0.$$

An empirically accepted triple with $\widehat{\Phi}_{t,k}(x;\lambda) \ge -\varepsilon_{\mathrm{risk},t}$ satisfies, on $\mathcal{E}_\delta \cap \mathcal{M}_\zeta$,

$$r_t(x) \le (1-\lambda) d_{t,k} + e_t(x) + \omega_{t,k}(x;\lambda). \tag{41}$$

Alternatively, the strengthened gate $\widehat{\Phi}_{t,k}(x;\lambda) - \omega_{t,k}(x;\lambda) \ge -\varepsilon_{\mathrm{risk},t}$ preserves the unmodified main certificate. Both claims have probability at least $1-\delta-\zeta$.

Proof. On $\mathcal{M}_\zeta$, empirical acceptance implies $\Phi_{t,k}(x;\lambda) \ge -\varepsilon_{\mathrm{risk},t} - \omega_{t,k}(x;\lambda)$. Apply Proposition 2 with this enlarged tolerance. The strengthened gate instead implies the original exact gate. A union bound proves the probability statement without assuming independence between statistical and numerical error events. □

A.8.1 AN EXPLICIT OPTIONAL CERTIFICATE ON A FINITE GRID

The following construction shows how the numerical radius can be obtained under additional independent-scenario assumptions. It is conservative and is not claimed to be the radius used in the empirical implementation. Condition on the history, working coupling, witnesses, and a finite set $\mathcal{Q}$ of J triples (x,k,λ), all fixed before the validation scenarios are drawn. For each triple a, draw S independent scenarios from its specified working joint law and denote its scalar margins by $Y_{a,s}$, with generic law Y_a. Samples may be shared across triples; only the sequence for each fixed triple must be i.i.d. A separate validation sample can be used if the triples were constructed with exploratory samples.

An integrable-risk perturbation inequality. For any integrable Y,

$$\mathrm{LCVaR}_\alpha(Y) = \sup_{v\in\mathbb{R}} \left\{ v - \frac{1}{\alpha}\mathbb{E}[(v-Y)_+] \right\}. \tag{42}$$

To verify this identity, the right and left derivatives of the concave expression in braces are $1 - F_Y(v)/\alpha$ and $1 - F_Y(v^-)/\alpha$. A maximizer is an α-quantile, with $F_Y(v^-) \le \alpha \le F_Y(v)$. Its value is the average of the lowest α probability mass, including any required fraction of the atom at v, which is exactly (20). This is the lower-tail version of the CVaR variational representation (Rockafellar & Uryasev, 2002). For two integrable variables under any coupling, the pointwise inequality $|(v-Y)_+ - (v-Y')_+| \le |Y-Y'|$ therefore gives

$$|\,\mathrm{LCVaR}_\alpha(Y) - \mathrm{LCVaR}_\alpha(Y')| \le \alpha^{-1}\mathbb{E}|Y-Y'|. \tag{43}$$

A finite-sample radius. Choose $L > 0$ before the validation sample and clip a scalar to $Y^L = \max\{-L, \min\{Y, L\}\}$. Define

$$h_S = \sqrt{\frac{\log(2J/\zeta)}{2S}}, \qquad \widehat{T}_a = \frac{1}{S}\sum_{s=1}^{S} |Y_{a,s} - Y^L_{a,s}|.$$

For triple $a = (x,k,\lambda)$, write $v_{a,j} = \mu^z_{t-1,j}(x) - b_{t,k,j}(\lambda)$ and $s_{a,j} = \sigma^z_{t-1,j}(x)$. For a normal variable of mean v and standard deviation s, let

$$H_L(v,s) = \begin{cases} s\varphi((L-v)/s) + (v-L)[1-\Phi_{\mathrm{N}}((L-v)/s)], & s > 0, \\ [v-L]_+, & s = 0. \end{cases}$$

Direct normal integration gives $H_L(v,s) = \mathbb{E}[(V-L)_+]$. Since Y_a is a minimum of coordinate margins,

$$\mathbb{E}|Y_a - Y_a^L| \le T_a := \sum_{j=1}^{M} [H_L(v_{a,j}, s_{a,j}) + H_L(-v_{a,j}, s_{a,j})]. \tag{44}$$

Indeed, $(-L - Y_a)_+$ is the maximum of the coordinate lower-tail excesses and is bounded by their sum; $(Y_a - L)_+$ is bounded by the sum of the upper-tail excesses. This argument does not require independent coordinates.

Proposition 4 (Uniform finite-grid empirical-risk bound)**.** *Under the independent-scenario conditions above, with conditional probability at least $1-\zeta$, simultaneously for all $a \in \mathcal{Q}$,*

$$|\widehat{\Phi}_a - \Phi_a| \le \omega_a := \frac{\widehat{T}_a + T_a + 2L\min\{1, h_S\}}{\alpha}. \tag{45}$$

Here $\widehat{\Phi}_a$ is the empirical risk in (40) *of the original, untruncated scenarios.*

Proof. For each triple, apply the Dvoretzky–Kiefer–Wolfowitz–Massart inequality (Massart, 1990) to the clipped i.i.d. samples. A union bound over J triples gives, with conditional probability at least $1-\zeta$,

$$\sup_y |\widehat{F}_a^L(y) - F_a^L(y)| \le \min\{1, h_S\} \quad \text{for all } a.$$

For variables supported in $[-L, L]$, a maximizer in (42) may be restricted to $[-L, L]$ and $\mathbb{E}[(v - Y^L)_+] = \int_{-L}^{v} F_a^L(y)\,dy$. The uniform CDF bound therefore changes every variational objective by at most $2L\min\{1, h_S\}/\alpha$, and hence bounds the difference between the clipped empirical and clipped population risks by this amount. Equation (43), applied to the empirical coupling of $Y_{a,s}$ with $Y_{a,s}^L$, adds $\widehat{T}_a/\alpha$. Applied to the population clipping coupling, it adds at most T_a/α by (44). The triangle inequality proves the result. □

Uniformity permits selection of the best triple after examining the validation scores. A pointwise interval at a preselected triple is not interchangeable with such a uniform guarantee. The construction does not justify arbitrary adaptive queries outside $\mathcal{Q}$, non-i.i.d. scenario-generation schemes, or an unreported approximation to the empirical risk itself. A grid-selected level is sufficient for the main certificate; maximality over a continuous interval is unnecessary.

Using numerically certified repair in the LLM comparison. If the selector uses empirical accepted levels $\widehat{\lambda}(x)$, let $\omega(x)$ be a valid one-sided error bound for that triple in Proposition 3. On the joint error event, define the corrected tables

$$V_{t,k}^{\mathrm{num}}(x) = \widehat{\lambda}(x) d_{t,k} - e_t(x) - \omega(x), \qquad \mathcal{B}_{t,k}^{\mathrm{num}}(x) = d_{t,k} - V_{t,k}^{\mathrm{num}}(x).$$

The corrected certificate $\mathcal{B}_{t,k}^{\mathrm{num}}$ is nonnegative and bounds $r_t(x)$ at all candidates eligible for that comparison. Conditional on the fixed validated table, the finite-sum proof of Lemma 5 applies unchanged, with $V_{t,k}^{\mathrm{num}}$ in the covariance and its oscillation in the remainder constant. This yields an expected-certificate improvement only when alignment and the step condition hold for the *corrected* certificate, not merely for an uncorrected empirical score. If the strengthened gate is used, the original slack $e_t(x)$ suffices, with the accepted tested level in place of an unverified exact maximizer. All compared tables, including the card score, are frozen before varying the reweighting strength. Uniform numerical validation, as in Proposition 4, permits subsequent selection among those validated triples without a separate union bound over the reweighting choices. It does not imply the corresponding result for an unvalidated continuously optimized level.

A.9 Final selection, exploration, and path aggregation

Candidate-wise certificates and distributional comparisons. The exact-risk certificate is candidate-wise and uniform on the same GP confidence event. Card generation, KL-constrained reweighting, path mixing, and diversity processing may be arbitrary for this pointwise result, provided the final candidate retains at least one valid accepted path certificate. No exponential-family representation of the final selection law is required for (9)–(11). If several paths certify a candidate,

the smallest of their valid bounds may be used. In contrast, the expected-certificate improvement in (12) requires the specified reweighting law, or one of the explicit extensions below. Pointwise validity does not itself imply that an arbitrary final sampling or maximization rule inherits a distributional improvement.

An exploration candidate without a path certificate still has the generic confidence bound

$$r_t(x) \le H_t(x) := \max_{x' \in \mathcal{X}} u_t(x') - \ell_t(x). \tag{46}$$

This follows from $g^* \le \max u_t$ and $g(x) \ge \ell_t(x)$. A fallback that actually evaluates x_t^0 receives B_t^0; this is a sufficient option, not a claim about unprovided code.

A fixed independent exploration component. Let s be any fixed exploration distribution on $\mathcal{X}$, possibly assigning mass outside the certified support $\mathcal{S}$, and let $q_\eta = (1-\rho)p_\eta + \rho s$ for fixed $0 \le \rho \le 1$. Define a branch-wise expected certificate by

$$\mathfrak{B}_\eta = (1-\rho)\mathbb{E}_{p_\eta}\mathcal{B}_{t,k} + \rho \mathbb{E}_s H_t.$$

On $\mathcal{E}_\delta$, applying the appropriate pointwise bound in each branch and averaging gives $\mathbb{E}_{q_\eta} r_t \le \mathfrak{B}_\eta$. When $a_{t,k} > 0$ and the main step condition holds,

$$\begin{aligned} \mathfrak{B}_\eta - \mathfrak{B}_0 &= (1-\rho)\big[\mathbb{E}_{p_\eta}\mathcal{B}_{t,k} - \mathbb{E}_{p_0}\mathcal{B}_{t,k}\big] \\ &\le -(1-\rho)\eta a_{t,k}/2. \end{aligned} \tag{47}$$

Thus a common exploration component scales the certificate improvement by its guided fraction. It does not erase it when $\rho < 1$. The comparison uses the same s, ρ, and generic certificate in both laws; it does not certify physical safety of exploration or compare true regrets.

Fixed mixtures of witness paths. For each path k, form a law $q_{\eta,k}$ and its branch-wise expected certificate $\mathfrak{B}_{\eta,k}$ as above. For fixed nonnegative path weights π_k summing to one, actual sampling from $\sum_k \pi_k q_{\eta,k}$ has expected regret bounded by $\sum_k \pi_k \mathfrak{B}_{\eta,k}$ on the confidence event. Certificate changes add with exactly the same weights. In particular, if each path satisfies its own positive alignment and accepted-step conditions, the total reduction is at least $\sum_k \pi_k(1-\rho_k)\eta_k a_{t,k}/2$. Alternatively, signed per-path bounds follow directly from (32). This comparison preserves the path weights, supports, and certificate functions. It does not identify a mixture of separately normalized exponential tilts with a single tilt of an arbitrarily chosen global score. Reoptimizing path weights or rebuilding supports requires an additional analysis.

Diversity processing and deterministic choices. A deterministic argmax of p_η need not have the expected-certificate improvement of a sample from p_η. Nor is an arbitrary diversity replacement guaranteed to preserve that improvement. For a common postprocessor $K(y \mid x)$ and a common output certificate $B(y)$, the appropriate pre-processing utility is $KB(x) = \sum_y K(y \mid x)B(y)$; the alignment test must then concern its negative rather than silently reusing $V_{t,k}(x)$. This follows directly from $\mathbb{E}_{p_\eta K} B = \mathbb{E}_{p_\eta} KB$. Whole-batch processing similarly requires a joint proposal-law analysis. Another sufficient route is to apply the fixed-history comparison at each slot conditional on earlier selected batch members, provided that slot's actual law, available support, and scores satisfy the stated assumptions. No new real labels are required between such within-batch selections.

Witness retention and scope. For synchronous batches, every final member can be certified with the same pre-batch posterior. This requires no fictitious within-batch mean update. The bounds hold jointly on $\mathcal{E}_\delta$ and on the numerical error event when needed. If a certified candidate is retained in the post-evaluation frontier, its true utility also bounds the best retained utility from below. Without retention, a certificate for that evaluated candidate is not a guarantee about the next retained frontier. Under noisy observations, retaining the observed-best point alone does not ensure that the true-best point is retained.

The new language-guidance result proves a specific additional fact: positive alignment with net certified repair, under the stated finite step, reduces the expected path certificate at a fixed history. It does not assume or prove that every sampled candidate has positive net repair. Neither a smaller expected certificate nor $B_t^{\mathrm{in}} < B_t^0$ orders the true regrets of two different selectors, and the displayed GP-UCB width need not be the reference's tightest available certificate. The theorem does not

establish that a particular language model supplies positive alignment or that a suitable witness will be found. These remain algorithmic or empirical questions. No cumulative-regret rate, first-hit acceleration rate, or universal superiority follows from this one-step analysis alone.

B Additional Implementation Details

The following definitions specify the decision-layer computations summarized in the main text. They do not replace the implementation audit needed for the hybrid empirical runs.

B.1 Card consensus, adherence, and repair utility

When several independent LLM proposals describe the same witness, variable direction, and objective role, we merge them into one card. Let $\bar{\gamma}_{t,k,c}$ be its mean proposal confidence and $D_{t,k,c}$ the disagreement score formed from confidence dispersion and missing support. The validated card score is

$$S_{t,k,c} = \bar{\gamma}_{t,k,c} \exp(-\lambda_{\mathrm{dis}} D_{t,k,c}). \tag{48}$$

This score controls the initial card weight but never changes the GP training data. For a candidate x and witness $w_{t,k}$, each directional rule induces a nonnegative violation. For example, an "increase" rule and a "decrease" rule with desired change Δ_j give

$$L^{\uparrow}_{c,j}(x; w_{t,k}) = \max\{w_{t,k,j} + \Delta_j - x_j, 0\}, \qquad L^{\downarrow}_{c,j}(x; w_{t,k}) = \max\{x_j - (w_{t,k,j} - \Delta_j), 0\}. \tag{49}$$

Interval, hold, and category-switch rules use the corresponding distance to their legal interval or reference value. Writing the resulting aggregate as $L_{t,k,c}(x; w_{t,k}) \geq 0$, the candidate adherence is

$$A_{t,k,c}(x \mid w_{t,k}) = \exp\left[-L_{t,k,c}(x; w_{t,k})/T_A\right]. \tag{50}$$

The card is then screened against the posterior rather than trusted from its language confidence. Let $z^{(s)}_{t,j}(x)$ be a sampled normalized margin and $b_{t,k,c,j}(w_{t,k})$ the card-specific continuation requirement. Define

$$Z^{(s)}_{t,k,c}(x) = \min_j \left[z^{(s)}_{t,j}(x) - b_{t,k,c,j}(w_{t,k}) \right], \quad R_{t,k,c}(x) = \mathrm{LCVaR}_{\alpha}\left(Z_{t,k,c}(x) \mid \mathcal{D}_{t-1}\right). \tag{51}$$

The posterior-screened repair utility is

$$U_{t,k,c}(x) = \sigma\left(R_{t,k,c}(x)/T_R\right) A_{t,k,c}(x \mid w_{t,k}), \qquad U^{\mathrm{card}}_{t,k}(x) = \sum_{c \in \mathcal{H}_{t,k}} \omega_{t,k,c} U_{t,k,c}(x). \tag{52}$$

Thus a candidate receives high card utility only when it both follows the semantic direction and has a robust posterior prospect of repairing the weakest target. The semantic variable names used in this step are mapped to legal feature columns before scoring; the LLM is never required to interpret latent fingerprint coordinates.

B.2 Real-evaluation feedback

After evaluating a candidate, BRIDGE-MO compares its joint deficit with that of the owning witness. Let $d_t(w) = \max\{-b(w), 0\}$ and let $\mathcal{P}_{t,k}$ be the protected-objective set of witness k. We use the bounded progress and protection terms

$$g_t(x; w) = \frac{d_t(w) - d_t(x)}{d_t(w) + 0.5}, \qquad v_t(x; w) = \sum_{j \in \mathcal{P}_{t,k}} \max\{-z_j(x), 0\}, \tag{53}$$

and assign the observed intervention reward

$$r_t(x; w) = \mathrm{clip}_{[-1,1]}\left[\tanh\left(g_t(x; w) - \lambda_P v_t(x; w)\right)\right]. \tag{54}$$

A candidate can therefore receive a positive reward only when it reduces the joint deficit without excessively damaging protected objectives. If several cards explain the same candidate, responsibility is assigned by

$$q_{t,i,c} = \frac{A_{t,k,c}(x_i)\omega_{t,k,c}}{\sum_{c'} A_{t,k,c'}(x_i)\omega_{t,k,c'}}, \qquad \log \omega_{t+1,k,c} = \log \omega_{t,k,c} + \eta_{\omega}\, \bar{r}_{t,k,c}, \tag{55}$$

where $\bar{r}_{t,k,c}$ is the responsibility-weighted observed reward. Weights have a positive floor; repeated failures trigger card refresh rather than allowing one noisy evaluation to erase a hypothesis permanently. These updates are based only on observed objective values and are separate from GP posterior fitting.

C EXECUTABLE BRIDGE-MO SPECIFICATION

This section specifies the final decision rules used by the archived experiments. Each candidate is a row of a frozen finite pool; its row identifier is retained and a row is never evaluated twice. Continuous inputs are affine-scaled to $[0, 1]$ using pool bounds before fitting the surrogate or computing distances. Raw outputs are retained for reporting.

C.1 TARGET MARGINS AND EVALUATION PROTOCOL

For objective j, let $d_j = 1$ for maximization and $d_j = -1$ for minimization, let τ_j be the fixed target, and let $s_j > 0$ be the fixed task scale. The normalized margin and joint margin are

$$z_j(x) = d_j \frac{f_j(x) - \tau_j}{s_j}, \qquad Q(x) = \min_{1 \le j \le M} z_j(x). \tag{56}$$

Joint attainment is exactly $Q(x) \ge 0$. The primary comparison uses five shared initial rows, total budget $T = 100$, batch size $q = 4$, and no replacement. The five rows are evaluated before any method-specific acquisition and count toward T. The module and model ablations use a separate frozen QM9 five-objective protocol with the same five shared initial evaluations, budget, and batch size. In that protocol LLM warm-start denotes only the preference applied after the common initial evaluations; it does not add free evaluations.

C.2 WITNESS FRONTIER AND PATH OWNERSHIP

The primary witness is

$$w_{t,1} = \arg\max_{x \in \mathcal{D}_t} Q(x), \tag{57}$$

with ties broken by the smallest pool row identifier. Auxiliary witnesses are eligible when their Q is above the larger of the $(1 - 0.35)$ archive quantile and the primary quality minus 0.90, with Pareto and near-feasible points also eligible. The separation score is

$$\delta(x, x') = 0.45 \frac{\|u(x) - u(x')\|_2}{\sqrt{D}} + 0.45 \frac{\|\tanh z(x) - \tanh z(x')\|_2}{\sqrt{M}} + 0.10\, \overline{\mathbf{1}[r(x) \ne r(x')]}, \tag{58}$$

where u is the normalized input and r is the target-role vector. Auxiliary points are added by farthest-first maximization of their minimum separation. The frontier starts with two witnesses and grows to at most five at evaluations 20, 45, and 75.

Each witness has independent cards, path state, and evaluation count. With path logit $a_{t,k}$, the primary path receives a prior bonus $\log(1.5)$:

$$\widetilde{\pi}_{t,k} = \mathrm{softmax}_k \left(a_{t,k} + \mathbf{1}[k = 1] \log(1.5) \right), \qquad \pi_{t,k} = \max \left(\widetilde{\pi}_{t,k}, \min(0.08, 0.95/K_t) \right), \tag{59}$$

followed by renormalization. Real-evaluation feedback updates only the owning path, $a_{t+1,k} = \mathrm{clip}_{[-3,3]}(a_{t,k} + 0.20 \widehat{r}_{t,k})$. Thus an auxiliary route cannot be removed by a single noisy observation.

C.3 CONTINUATION RISK AND BASE CANDIDATE DISTRIBUTION

Each objective surrogate supplies a posterior mean and standard deviation. Joint posterior draws use the fitted residual correlation model; the final configuration uses 512 draws. For witness k and continuation level $\lambda \in [0, 1]$,

$$b_{t,k,j}(\lambda) = (1 - \lambda) \min\{z_j(w_{t,k}), 0\}, \qquad R_{t,k}(x; \lambda) = \mathrm{LCVaR}_{0.30} \left[\min_j \{ z_{t,j}^{(s)}(x) - b_{t,k,j}(\lambda) \} \right]. \tag{60}$$

The algorithm scans 21 equally spaced values and chooses the largest reachable value whose lower-tail risk is nonnegative for an eligible candidate.

Before the first joint hit, the base score is

$$s_0 = 0.55\ \mathrm{std}(\log \widehat{P}) + 0.85\ \mathrm{std}(\widehat{Q}) + 0.45\ \mathrm{std}(\widehat{m}), \qquad p_0 = \mathrm{softmax}(s_0 / 0.18), \tag{61}$$

where $\widehat{P}$, $\widehat{Q}$, and $\widehat{m}$ are posterior joint-hit probability, expected joint margin, and expected bottleneck margin. After a joint hit, the score is

$$s_0 = 0.10\ \mathrm{std}(\log \widehat{P}) + 0.20\ \mathrm{std}(\widehat{Q}) + 0.10\ \mathrm{std}(\widehat{m}) + 1.00\ \mathrm{std}(\widehat{F}), \qquad p_0 = \mathrm{softmax}(s_0/0.18), \tag{62}$$

where $\widehat{F}$ is distance-weighted unexplored-frontier potential.

C.4 Card tilt, safety, and batch selection

For each witness, the accepted-card utility is the posterior-screened $U^{\mathrm{card}}_{t,k}$ defined in Eqs. (51)–(52). The card distribution is

$$p^{\mathrm{card}}(x;\eta) \propto p_0(x)\exp\{\eta\ \mathrm{std}(U^{\mathrm{card}}(x))\}, \quad \eta \le 100, \quad \mathrm{KL}(p^{\mathrm{card}}\|p_0) \le 0.05. \tag{63}$$

The largest feasible η is used; minimum intervention strength is 0.15 before attainment and 0.05 after attainment. Independent safety exploration uses

$$p^{\mathrm{safe}}(x) = \mathrm{softmax}\left([\mathrm{std}(\mu_Q + 2\sigma_Q) + 0.35\,\mathrm{std}(\sigma_Q)]/0.18\right), \qquad p = (1-\rho)p^{\mathrm{card}} + \rho p^{\mathrm{safe}}, \tag{64}$$

where $\rho \in [0.08, 0.30]$, or $\rho = 0.55$ if no continuation is reachable.

The shortlist contains 50% optimistic, 20% witness-local, 15% novel, and 15% uniform candidates. Its sizes are NMD18 3,000, PFD 1,052, QM9 2,500, Sandwich 2,500, and 5,000 for each Tartarus task. Within a batch, after the first path slots, choose the row maximizing

$$\log p(x) + 0.55 \log\left(1 + d_{\min}(x, B)/0.20\right), \tag{65}$$

without replacement. Path slots are ordered by $\pi_{t,k}/\sqrt{1+n_{t,k}}$. The tie-break is always the smallest pool row identifier.

C.5 Complete executable round

1. Load the frozen pool, directions, targets, scales, representation, seed, and the five shared initial row identifiers; evaluate them and set $t = 5$.
2. Fit one surrogate per objective, fit residual correlations, draw 512 joint posterior samples, and compute margins, bottlenecks, Pareto status, and archive diagnostics.
3. Construct the witness frontier, keeping independent cards and paths.
4. Validate existing cards; when a card is missing or expired, query the LLM for that witness only. The prompt contains the task schema, witness margins, and at most 20 recent post-card observations, but no candidate label that can be selected directly.
5. Compute continuation risk, path weights, p_0, card tilt, safety mix, shortlist, and the diverse batch rule above.
6. Evaluate $q = 4$ selected rows with the true oracle, append raw outputs, and refit the surrogate.
7. Compute the real progress/protection reward from Eqs. (53)–(54); update only the owning path and cards. Refresh after three consecutive failures or 24 evaluations, with a minimum refresh interval of 16.
8. Increase t by q and repeat until $t \ge T$. Record every row, source, witness, card responsibility, posterior diagnostic, best margin, first-hit index, and hypervolume.

D Task and Baseline Details

The tables below complement the dataset and comparator introductions in the main experimental section.

D.1 Primary and Ablation Experimental Protocols

The primary comparison and the component studies are deliberately separate protocols. The primary comparison uses the six task definitions listed below; its QM9 task has eleven simultaneous

objectives and its reported main-table values are computed from the matched main-result package. The module-ablation and language-model-sensitivity studies use a separate, frozen five-objective QM9 protocol with the same candidate-pool convention, five shared initial evaluations, 100 total evaluations, and batch size $q = 4$. Thus, the first-hit values 15.4 ± 3.4 in the primary QM9 table and 12.2 in the ablation table are not estimates of the same random variable: they come from different objective sets and different result packages. The ablation is a within- protocol causal comparison, not a replacement for the eleven-objective main result.

In both protocols, the five initial candidates are evaluated before any method-specific acquisition is allowed to act and count toward the budget. The phrase *LLM warm start* refers only to the early post-initialization acquisition stage: it changes the preference over unevaluated candidates using validated intervention cards, but it does not add candidates to, or alter, the shared initial set. The "w/o LLM warm-start" row therefore keeps the same five initial evaluations and removes only this early language-guided preference; cards, posterior fitting, safety screening, and batch selection are changed only in the ablation explicitly naming them.

D.2 Complete task definitions

Table 2 records the exact pool size, input dimension, representation, and shortlist used in the primary result package. The ranges are raw data ranges or fixed design ranges; the surrogate receives the pool-bound normalized values.

Table 2: Task manifest for the six primary finite-pool benchmarks.

Task	Pool	Dim.	Input representation	Shortlist and protocol
NMD18	3,000	15	Composition, crystal statistics, density, space-group, and pair-distance descriptors	Frozen finite pool; shortlist 3,000.
PFD	1,052	7	Polymer molecular-weight, frontier-orbital, and band-gap descriptors	Frozen finite pool; shortlist 1,052.
QM9	20,000	12	Molecular size, atom counts, H-bonding, rings, flexibility, TPSA, logP, and valence	Frozen finite pool; shortlist 2,500; 11-objective primary task.
Sandwich	16,384	48	Ingredient, composition, process, assembly, sensory, and packaging variables	Deterministic pool, scramble seed 20260902; shortlist 2,500.
Tartarus emitter	386,980	13	GDB13 molecular descriptors and atom-fraction features	Frozen finite pool; shortlist 5,000.
Tartarus reaction	60,828	29	13 semantic descriptors plus 16 Morgan-SVD coordinates	Frozen finite pool; shortlist 5,000.

NMD18. The inputs are fraction Al, fraction Ga, fraction In, atom count, species count, mean and standard deviation of atomic number, cell volume, cell anisotropy, number density, space-group number, pair-distance mean and standard deviation, and nearest-neighbor mean and standard deviation. The objectives are formation energy (minimize, target 0.0228 eV atom^{-1}, scale 0.041834916) and band-gap absolute error (minimize, target 0.18248 eV, scale 0.572576026). Raw ranges are composition fractions $[0, 0.4]$, atom count $[10, 80]$, species count $[2, 4]$, mean atomic number $[10, 24.4]$, standard deviation $[2.4495, 20.0858]$, cell volume $[93.2344, 1137.3528]$ A^3, anisotropy $[1, 4.538]$, number density $[0.06815, 0.11397]$ atoms A^{-3}, space-group number $[12, 227]$, pair-distance mean $[3.6597, 7.7435]$ A, pair-distance standard deviation $[1.1158, 3.6174]$ A, nearest-neighbor mean $[1.804, 2.1685]$ A, and nearest-neighbor standard deviation $[0.0000574, 0.157479]$ A.

PFD. The inputs are molecular weight, number-average molecular weight, PDI, monomer mass, HOMO, LUMO, and band gap. The objectives are power conversion efficiency (maximize, target 7.0335 percent, scale 2.377896), open-circuit voltage (maximize, target 0.92 V, scale 0.134792), short-circuit current density (maximize, target 13.4 mA cm^{-2}, scale 4.024562), and fill factor (maximize, target 0.66635, scale 0.110390). Raw ranges are molecular weight $[3.9, 2535]$ kg mol^{-1}, number-average molecular weight $[0, 780]$, PDI $[1.05, 25]$, monomer mass $[196, 3144]$ g mol^{-1}, HOMO magnitude $[4.24, 6.18]$ eV, LUMO magnitude $[2.3, 4.9]$ eV, and band gap $[1.0, 2.39]$ eV.

QM9 primary task. The twelve inputs are molecular weight, heavy-atom count, heteroatom count, hydrogen-bond donors, hydrogen-bond acceptors, ring count, aromatic-ring count, rotatable-bond count, fraction sp^3, TPSA, logP, and valence electrons. Their raw ranges are $[27.026, 152.037]$ g mol^{-1}, $[2, 9]$, $[0, 7]$, $[0, 5]$, $[0, 7]$, $[0, 8]$, $[0, 2]$, $[0, 6]$, $[0, 1]$, $[0, 123.47]$ A^2, $[-4.8966, 3.6128]$, and $[10, 56]$, respectively. The eleven fixed targets, in audit order, are electronic gap maximize 0.34 Hartree ($s = 0.047704876$), polarizability maximize 95.0 Bohr^3 ($s = 8.162799$), dipole moment minimize 1.2 Debye ($s = 1.546847$), HOMO energy minimize -0.26042 Hartree ($s = 0.022126992$), LUMO energy minimize 0.08702 Hartree ($s = 0.046655415$), electronic spatial extent minimize 2211.43422 Bohr^2 ($s = 280.486033$), zero-point vibrational energy minimize 0.272737 Hartree ($s = 0.033302216$), heat capacity maximize 39.1756 cal mol^{-1} K^{-1} ($s = 4.061521$), atomization internal energy minimize -2467.355975 kcal mol^{-1} ($s = 239.229758$), rotational constant A maximize 2.206046 GHz ($s = 1.895796$), and rotational constant B maximize 0.599632 GHz ($s = 0.570843$).

Sandwich oracle. The 48 inputs are grouped into bread, protein, dairy/fat, produce, sauce/seasoning, process, and assembly variables. Bread variables are bread mass $[45, 120]$ g, whole-grain, sourdough, and rye fractions $[0, 1]$, seed content $[0, 20]$ g, moisture $[28, 48]$ percent, and slice thickness $[6, 20]$ mm. Protein variables are protein mass $[20, 150]$ g, five composition fractions $[0, 1]$, lean fraction $[0.3, 1]$, and freshness $[0, 6]$ days. Dairy/fat variables are cheese $[0, 60]$ g, low-fat fraction $[0, 1]$, avocado $[0, 70]$ g, olive oil $[0, 15]$ g, mayonnaise $[0, 30]$ g, and butter $[0, 15]$ g. Produce variables are vegetables $[0, 180]$ g, five fractions $[0, 1]$, fruit $[0, 60]$ g, whole-fruit fraction $[0.5, 1]$, variety count $[1, 7]$, and freshness $[0, 7]$ days. Sauce variables are sauce $[0, 45]$ g, sugar $[0, 12]$ g, pH $[2.8, 6.5]$, viscosity $[0.2, 12]$ Pa s, sodium $[0, 1000]$ mg, herbs $[0, 8]$ g, and heat $[0, 50]$ kSHU. Process variables are toast temperature $[80, 240]$ C, toast time $[0, 12]$ min, press force $[0, 150]$ N, rest time $[0, 30]$ min, and serving temperature $[4, 75]$ C. Assembly variables are four scores in $[0, 1]$ and holding time $[0, 180]$ min.

The deterministic oracle uses $\mathrm{peak}(v; c, h) = \exp[-(v-c)^2/(2h^2)]$, $\mathrm{ramp}(v; l, u) = \mathrm{clip}((v-l)/(u-l), 0, 1)$, $\mathrm{adeq}(v; s, p) = p\,\mathrm{clip}(v/s, 0, 1)$, and $\mathrm{mod}(v; f, z, p) = p\,\mathrm{clip}((z-v)/(z-f), 0, 1)$. Energy, nutrient densities, saturated and unsaturated fat, sodium density, added-sugar percentage, and saturated-fat percentage are computed from the ingredient coefficients in the frozen oracle. HEI adequacy is $100/60$ times nine adequacy terms for fruit, whole fruit, vegetables, greens/beans, whole grains, dairy, protein, seafood/plant, and fatty-acid ratio. HEI moderation is $100/40$ times four moderation terms for refined grains, sodium, added sugar, and saturated fat. The six targets are HEI adequacy maximize 71.0 ($s = 5$), HEI moderation maximize 64.8 ($s = 5$), protein-fiber balance maximize 62.1 ($s = 5$), sensory quality maximize 48.2 ($s = 5$), structural integrity maximize 56.5 ($s = 5$), and estimated cost minimize 7.84 USD ($s = 1$).

Protein-fiber balance is the fourth root of protein balance, fiber balance, energy balance, and mass balance, with a bounded saturated-fat penalty. Sensory quality is the weighted bounded sum of freshness, toast, acid, salt, sauce balance, temperature, heat, savoriness, spread evenness, three pairing terms, and a bounded sinusoidal interaction. Structural integrity is the weighted bounded sum of leakage, viscosity, toast, press, slice geometry, compression, alignment, spread evenness, holding survival, and mass penalty. Cost is ingredient cost plus process cost plus quadratic waste. The final interaction is $r = 1.25 \sin(0.032\,\text{toast temperature} + 0.75\,\text{pH} + 0.018\,\text{press force})$. It adds $0.35r$ times whole-grain fraction to adequacy, subtracts $0.28r$ times processed-meat fraction from moderation, adds $0.55r$ times plant-plus-fish fraction to protein-fiber balance, adds r to sensory quality, and subtracts $0.65r$ from structural integrity, with each output clipped to its legal range. This is the executable six-output oracle used to construct the frozen pool.

Sandwich column names and exact coefficient map. The 48 pool columns, in order, are `bread_mass_g`, `whole_grain_fraction`, `sourdough_fraction`, `rye_fraction`, `seed_content_g`, `bread_moisture_pct`, `slice_thickness_mm`, `protein_mass_g`, `poultry_fraction`, `fish_fraction`, `plant_protein_fraction`, `egg_fraction`, `processed_meat_fraction`, `protein_lean_fraction`, `protein_freshness_days`, `cheese_mass_g`, `low_fat_dairy_fraction`, `avocado_mass_g`, `olive_oil_g`, `mayonnaise_g`, `butter_g`, `vegetable_mass_g`, `leafy_green_fraction`, `tomato_fraction`, `cruciferous_fraction`, `legume_fraction`, `pickled_fraction`,

`fruit_mass_g`, `whole_fruit_fraction`, `vegetable_variety_count`, `produce_freshness_days`, `sauce_mass_g`, `sauce_added_sugar_g`, `sauce_ph`, `sauce_viscosity_pa_s`, `added_sodium_mg`, `herb_spice_g`, `heat_level_kshu`, `toast_temperature_c`, `toast_time_min`, `press_force_n`, `rest_time_min`, `serving_temperature_c`, `layer_alignment`, `moisture_barrier`, `sauce_spread_evenness`, `packaging_compression`, and `holding_time_min`. Fractions within each recipe group are normalized after adding 0.04 to every supplied fraction.

For a compact executable notation, use b, w, d, r, e, m, h for the seven bread columns; $p, o, f, l, g, x, \ell, \phi$ for protein mass, poultry, fish, plant, egg, processed meat, lean fraction, and freshness; c, q, a, v, y, u for cheese, low-fat fraction, avocado, oil, mayonnaise, and butter; $V, L, T, C, G, K, F, H, n, \psi$ for vegetables, leafy greens, tomato, cruciferous, legume, pickled fraction, fruit, whole-fruit fraction, variety count, and freshness; S, A, P, ν, N, E, H_h for sauce, sugar, pH, viscosity, sodium, herbs, and heat; T_c, T_m, F_p, R_m, T_s for process; and A_l, B_m, E_s, C_p, H_o for assembly. The oracle first computes

$$E_b = b(2.62 - .24w - .08d), \tag{66}$$
$$E_p = 1.65o + 1.75f + 1.55l + 1.55g + 2.75x + .55(1 - \ell)(o + x), \tag{67}$$
$$E = \max\{120, E_b + pE_p + c(3.65 - .85q) + 1.6a + 8.84v + 6.8y + 7.2u \tag{68}$$
$$+ .32V + .60F + .85S + 4A + 5.5e\}, \tag{69}$$
$$\kappa = 1000/E, \quad X_F = F\kappa/150, \quad X_{WF} = FH\kappa/150, \quad X_V = V\kappa/85, \tag{70}$$
$$X_{GB} = [V(.72L + .82C + G)/85 + .60pl/85]\kappa, \tag{71}$$
$$X_{WG} = [b(.78w + .18r)/28.35 + e/56]\kappa, \quad X_D = c\kappa/42.5, \tag{72}$$
$$X_P = p\kappa/28.35, \qquad X_{SP} = [p(f + l)/28.35 + .35VG/28.35]\kappa. \tag{73}$$

Here pl means p times the plant fraction. The fat and sodium quantities are

$$F_{\text{sat}} = p(.018o + .022f + .010l + .032g + .095x) + c(.205 - .105q) + .12y + .51u + .022a, \tag{74}$$
$$F_{\text{unsat}} = p(.025o + .075f + .060l + .035g + .045x) + .145a + .92v + .55y + .42e, \tag{75}$$
$$R_F = F_{\text{unsat}}/\max(F_{\text{sat}}, .5), \quad G_R = b(1 - .82w)\kappa/28.35, \tag{76}$$
$$N_{\text{mg}} = N + 13.5px + 2.3p(o + f + l + g) + 13.8c + 10.5VK + 5S, \tag{77}$$
$$N_D = N_{\text{mg}}\kappa/1000, \quad A_S = 400A/E, \quad A_F = 900F_{\text{sat}}/E. \tag{78}$$

With $\text{peak}(v; c, h) = \exp[-(v - c)^2/(2h^2)]$, $\text{ramp}(v; l, u) = \text{clip}((v - l)/(u - l), 0, 1)$, $\text{adeq}(v; s, p) = p\,\text{clip}(v/s, 0, 1)$, $\text{mod}(v; f, z, p) = p\,\text{clip}((z - v)/(z - f), 0, 1)$, and $\text{ratio}(v; b_0, a_0, p_0) = p_0\,\text{clip}((v - b_0)/(a_0 - b_0), 0, 1)$, the first two outputs are

$$Y_A = \frac{100}{60}[\text{adeq}(X_F; .8, 5) + \text{adeq}(X_{WF}; .4, 5) + \text{adeq}(X_V; 1.1, 5) \tag{79}$$
$$+ \text{adeq}(X_{GB}; .2, 5) + \text{adeq}(X_{WG}; 1.5, 10) + \text{adeq}(X_D; 1.3, 10) \tag{80}$$
$$+ \text{adeq}(X_P; 2.5, 5) + \text{adeq}(X_{SP}; .8, 5) + \text{ratio}(R_F; 1.2, 2.5, 10)], \tag{81}$$
$$Y_M = \frac{100}{40}[\text{mod}(G_R; 1.8, 4.3, 10) + \text{mod}(N_D; 1.1, 2, 10) \tag{82}$$
$$+ \text{mod}(A_S; 6.5, 26, 10) + \text{mod}(A_F; 8, 16, 10)]. \tag{83}$$

The protein-fiber output uses

$$P_g = p(.29o + .25f + .17l + .13g + .22x) + .22c + .09b + .18e, \tag{84}$$
$$F_g = b(.018 + .075w) + .028V + .075VG + .024F + .080pl + .11e, \tag{85}$$
$$B_P = .45 + .55\,\text{ramp}(P_g\kappa; 42, 72), \quad B_F = .35 + .65\,\text{ramp}(F_g\kappa; 11, 24), \tag{86}$$
$$B_E = \text{peak}(E; 560, 150), \quad B_M = \text{peak}(b + p + c + a + V + F + S + v + y + u + e; 330, 95), \tag{87}$$
$$Y_P = 100\,\text{clip}((B_P B_F B_E B_M)^{1/4} - .12\,\text{clip}((A_F - 13)/9, 0, 1), 0, 1). \tag{88}$$

For sensory quality, define freshness $\exp(-.11\phi - .09\psi)$, toast $\text{peak}(T_c; 178, 32)\,\text{peak}(T_m; 5, 2.3)$, acid $\text{peak}(P; 4.25, .72)$, salt $\text{peak}(N_{\text{mg}}/M; 3, 1.1)$, sauce balance $\text{peak}(S/M; .075, .035)$, temperature $\text{peak}(T_s; 43, 17)$, and heat $\text{peak}(H_h; 13, 13)$, where $M = b + p + c + a + V +$

$F + S + v + y + u + e$. Savoriness is $\mathrm{clip}(.34o + .42f + .37l + .30g + .48x + .0045c + .026E_h, 0, 1)$, where E_h is herb mass. The remaining pairing terms are the square roots of the clipped products specified by their component fractions. The sensory coefficients, in order, are $(.10, .12, .10, .09, .10, .08, .06, .12, .08, .05, .04)$, followed by the bounded interaction $.02 \sin(3\,\mathrm{acid} + 2\,\mathrm{savoriness} + \mathrm{toast})$. Structural integrity uses wet load $.32V/180 + .25F/60 + .32S/45 + .20(m - 28)/20$, viscosity $\mathrm{peak}(\log(1 + \nu); \log(1 + 4.2), .55)$, press $\mathrm{peak}(F_p; 62, 36)$, slice geometry $\mathrm{peak}(h; 12, 3.5)$, compression $\mathrm{peak}(C_p; .28, .18)$, holding survival $\exp(-H_o/190)\exp(-R_m/65)$, and leakage $\mathrm{clip}(1 - \mathrm{wetload}(1 - .72B_m), 0, 1)$. Its coefficients, in order, are $(.18, .12, .12, .10, .10, .10, .12, .08, .08, -.06)$ for leakage, viscosity, toast, press, slice geometry, compression, alignment, spread, holding survival, and mass penalty.

Estimated cost is ingredient cost plus process cost and quadratic waste: the ingredient coefficients are .014 bread, .030/.052/.019/.025/.027 for poultry/fish/plant/egg/processed protein, .038 cheese, .030$avocado$,.065 oil, .022$mayonnaise$,.018 butter, .012$vegetables$,.017 fruit, .020$sauce, and$.028 seed; process cost is $.28 + .018T_m + .0009\max(T_c - 80, 0) + .004R_m + .0015H_o$; waste is $.008\max(M - 420, 0)^2/20$. Finally, $r = 1.25\sin(.032T_c + .75P + .018F_p)$ and the final corrections are $Y_A \leftarrow \mathrm{clip}(Y_A + .35rw, 0, 100)$, $Y_M \leftarrow \mathrm{clip}(Y_M - .28rx, 0, 100)$, $Y_P \leftarrow \mathrm{clip}(Y_P + .55r(l + f), 0, 100)$, $Y_S \leftarrow \mathrm{clip}(Y_S + r, 0, 100)$, and $Y_I \leftarrow \mathrm{clip}(Y_I - .65r, 0, 100)$. This coefficient map, together with the ranges above, is the complete deterministic Sandwich oracle.

Tartarus tasks. The emitter inputs are molecular weight, heavy-atom count, heteroatom count, hydrogen-bond donors, hydrogen-bond acceptors, ring count, aromatic-ring count, rotatable bonds, TPSA, logP, valence electrons, nitrogen fraction, and oxygen fraction. Their ranges are $[67.091, 226.347]$, $[5, 13]$, $[0, 6]$, $[0, 6]$, $[0, 6]$, $[1, 4]$, $[0, 4]$, $[0, 4]$, $[0, 125.34]$, $[-3.4765, 4.8205]$, $[26, 68]$, $[0, .384615]$, and $[0, .333333]$. The objectives are singlet-triplet gap minimize 0.697084629 eV ($s = 0.395698097$) and oscillator strength maximize 0.106296170 ($s = 0.148920059$).

The reaction inputs contain the same semantic descriptor family, with halogen fraction added, plus 16 Morgan-SVD coordinates used only by the surrogate. The 29 model dimensions are therefore 13 semantic descriptors and 16 latent coordinates. The five objectives are activation energy minimize 76.076603 kcal mol^{-1} ($s = 3.075751$), reaction energy minimize -12.089382 kcal mol^{-1} ($s = 4.506869$), synthetic accessibility minimize 7.064111 ($s = 0.389415$), molecular mass minimize 427.1528 g mol^{-1} ($s = 63.892494$), and lipophilicity error minimize 4.03895 ($s = 1.251084$). The latent coordinates are standardized frozen-pool features and are never assigned semantic names in an LLM prompt.

D.3 Final hyperparameters

Table 3: Core non-default settings of the final BRIDGE-MO configuration.

Component	Value	Meaning
Budget / initial / batch	$100/5/4$	Five shared evaluations, then batches of four.
Surrogate	GP; noise 10^{-5}	One objective model and residual correlation for joint draws.
Continuation	512 draws; $\alpha = .30$; 21 levels	Lower-tail route screening.
Witnesses	$K_0 = 2$, $K_{\max} = 5$; growth $20, 45, 75$	Quality plus farthest-first frontier.
Distance weights	$.45/.45/.10$	Input, margin, and role components.
Path controls	prior 1.5; floor .08; step .20	Independent route ownership.
Cards	max 10; refresh 24; minimum interval 16	Validated witness-specific interventions.
Reweighting	KL 0.05; $\eta_{\max} = 100$	Trust-region card tilt.
Safety / diversity	$\rho = .08$–$.30$; unreachable .55; coefficient .55	Exploration and batch separation.
LLM	temperature .25; timeout 420 s; five attempts	Schema-validated card generation.

Table 4: Scientific candidate pools and fixed-target objectives. "Min" and "Max" indicate the direction before normalization to the margin convention under the normalized target-margin convention introduced in Section 3.

Pool	Objectives	Targeted properties and provenance
NMD18	2	Formation energy (Min); absolute error to a 2.5-eV band-gap target (Min). NOMAD 2018 oxide pool (Sutton et al., 2019); material-representation context (Langer et al., 2022).
PFD	4	Power-conversion efficiency, open-circuit voltage, short-circuit current density, and fill factor (all Max). Polymer-photovoltaic measurements (Nagasawa et al., 2018).
QM9	11	Eleven molecular properties with task-specific fixed thresholds. Original molecular property pool (Ramakrishnan et al., 2014).
Sandwich	6	HEI adequacy, HEI moderation, protein–fiber balance, sensory quality, structural integrity (Max), and estimated cost (Min). Internal synthetic oracle; HEI-inspired structure (Shams-White et al., 2023).
Tartarus emitter	2	Singlet–triplet gap (Min) and oscillator strength (Max). GDB13 quantum-chemical benchmark (Nigam et al., 2023).
Tartarus reaction	5	Activation energy, reaction energy, synthetic accessibility score, molecular mass, and lipophilicity error (all Min). xTB-derived reaction benchmark (Nigam et al., 2023).

Table 5: Comparison methods and their role in the primary study. The label "MOEA" in Fig. 2 denotes MOEA/D-LLM.

Method	Role in the comparison
Hanaoka	Goal-oriented target-attainment probability for multi-objective inverse design (Hanaoka, 2021).
TAMO	Amortized in-context multi-objective optimization with a pretrained transformer policy (Zhang et al., 2026).
MOEA/D-LLM	LLM-assisted decomposition-based evolutionary optimization (Liu et al., 2023).
BOFormer	Non-Markovian reinforcement-learning policy for multi-objective BO (Hung et al., 2025).
MORBO	Trust-region multi-objective BO for high-dimensional search (Daulton et al., 2021b).
NSGA-II	Dominance and crowding-distance evolutionary baseline (Deb et al., 2002).
qParEGO	Randomized scalarization baseline for expensive multi-objective BO (Knowles, 2006).
qLogNEHVI	Numerically stabilized noisy expected-hypervolume-improvement baseline (Daulton et al., 2021a; Ament et al., 2023).

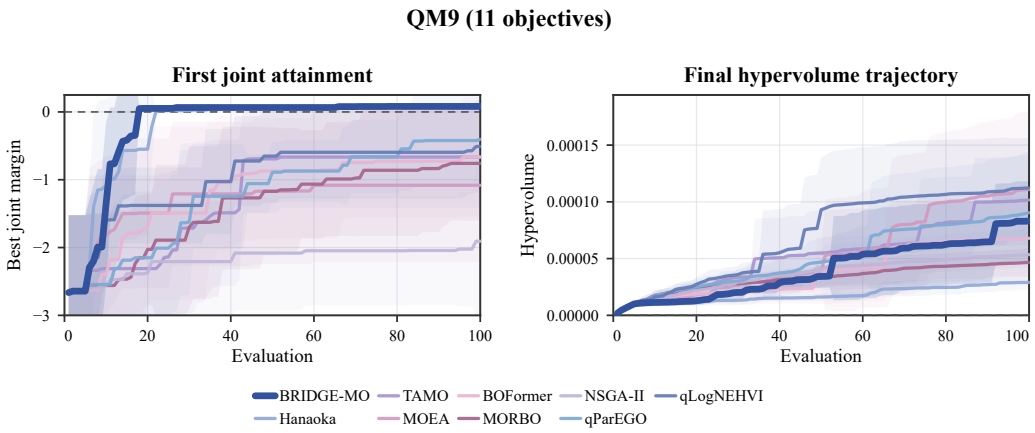

Figure 5: QM9 (11 objectives): first joint-attainment and hypervolume trajectories.

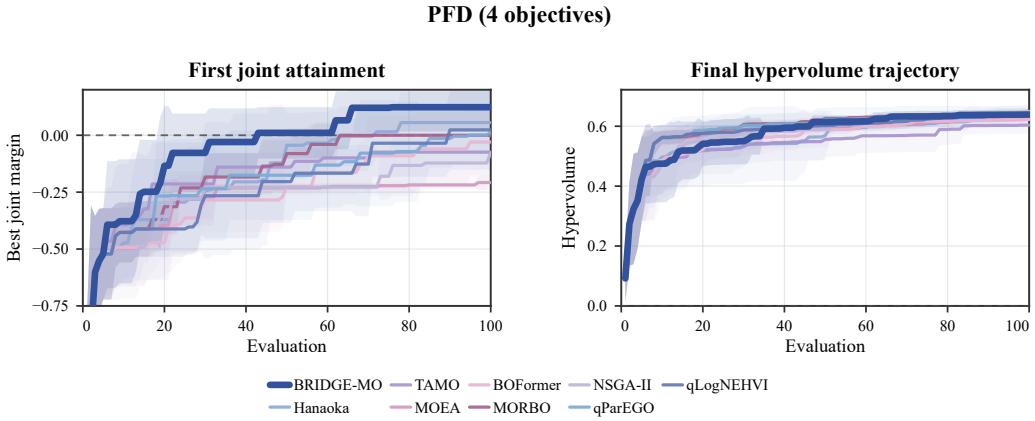

Figure 6: PFD (4 objectives): first joint-attainment and hypervolume trajectories.

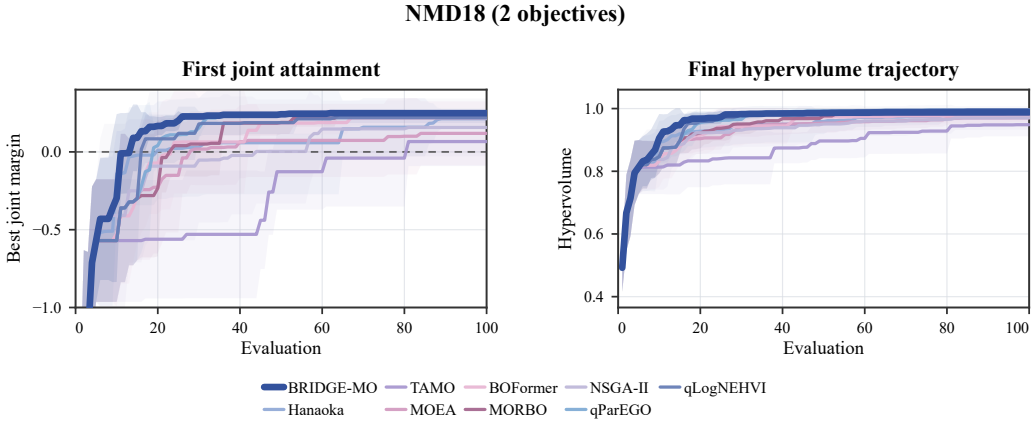

Figure 7: NMD18 (2 objectives): first joint-attainment and hypervolume trajectories.

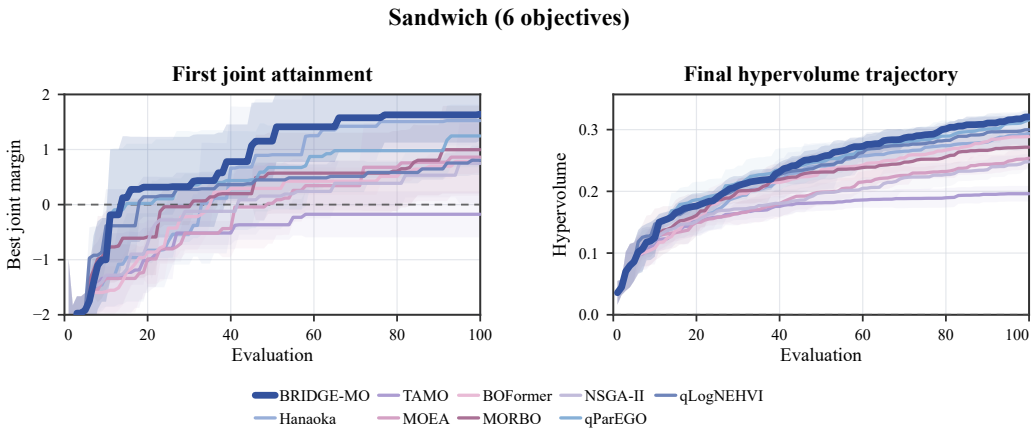

Figure 8: Sandwich (6 objectives): first joint-attainment and hypervolume trajectories.

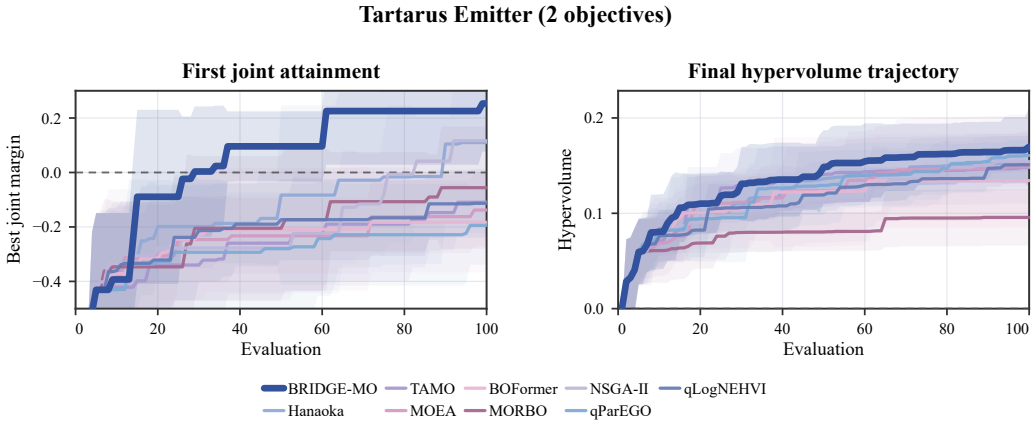

Figure 9: Tartarus emitter (2 objectives): first joint-attainment and hypervolume trajectories.

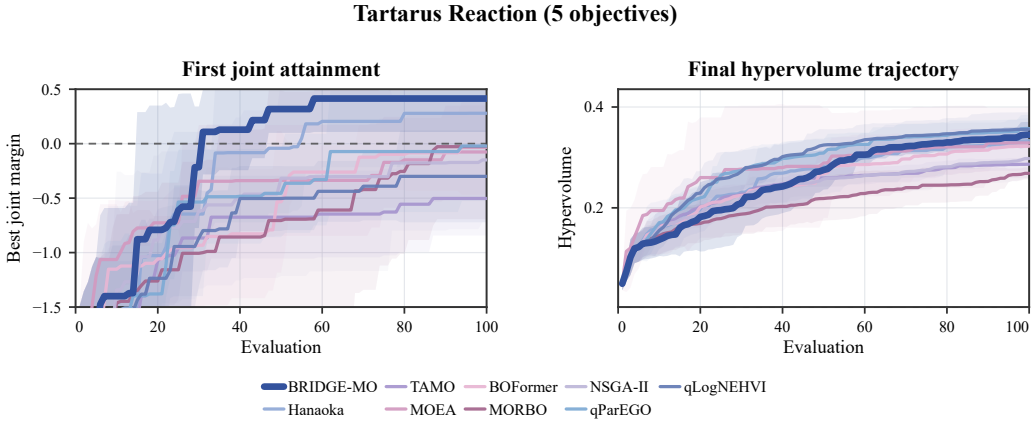

Figure 10: Tartarus reaction (5 objectives): first joint-attainment and hypervolume trajectories.

E Ablation and Model Sensitivity

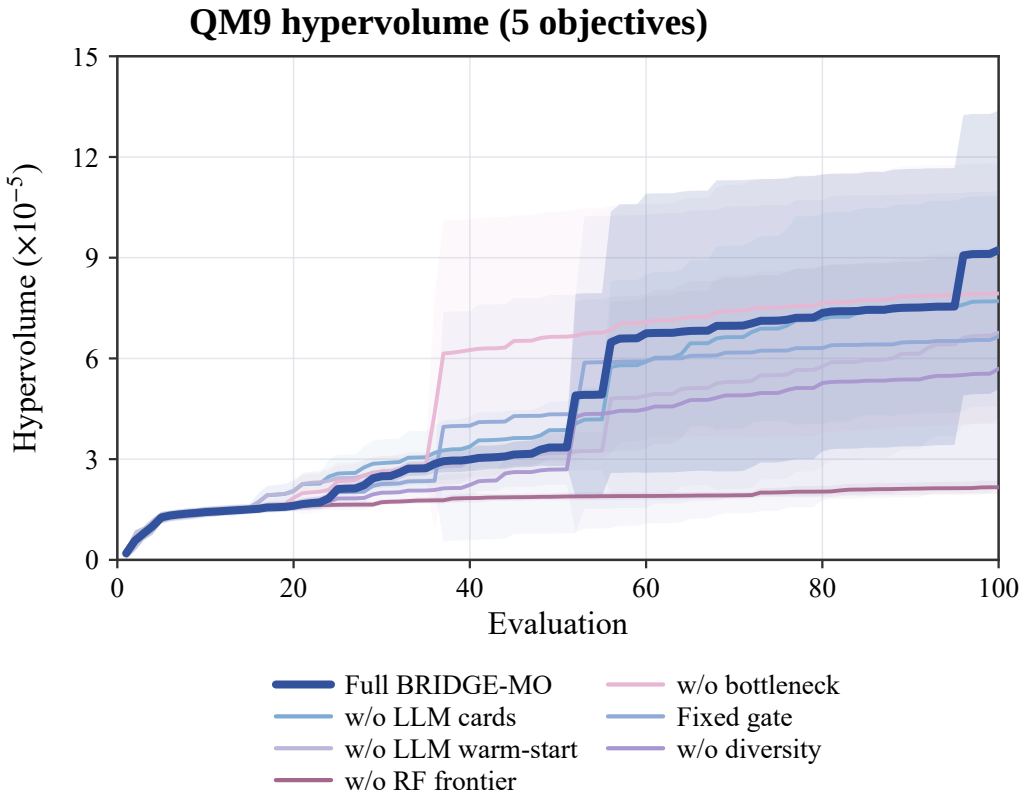

Figure 11: Hypervolume trajectories for the module ablation.

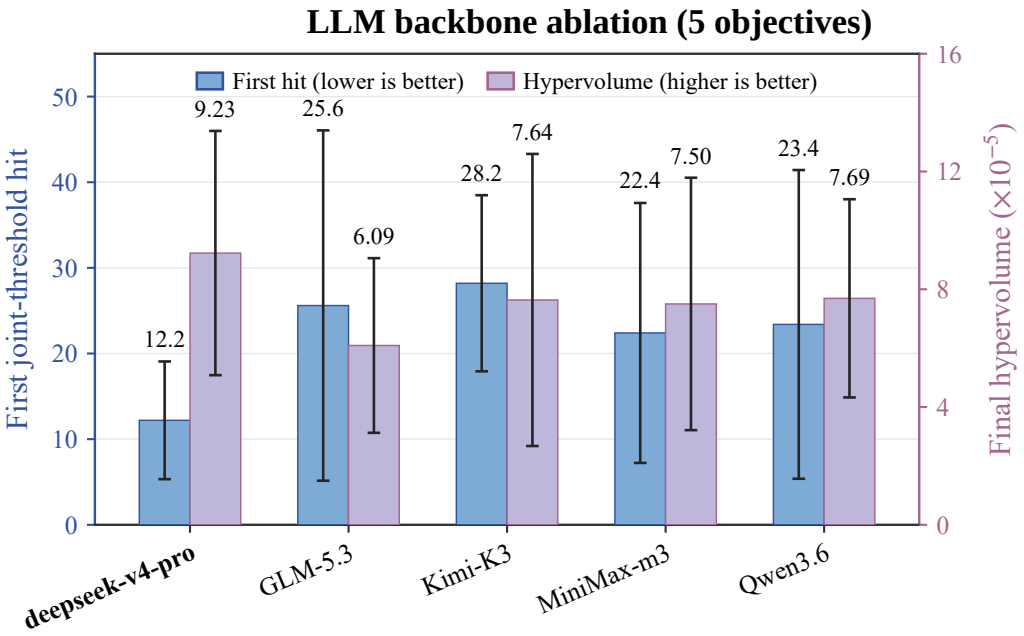

Figure 12: LLM model sensitivity with matched seeds and refresh schedule.

E.1 SELECTED ENDPOINT ABLATION VALUES

Table 6: Module ablation endpoint values on the matched QM9 five-objective package.

Variant	Final HV mean	HV std	First hit mean	Deficit AUC mean
Full model	9.23×10^{-5}	4.15×10^{-5}	12.2	0.230
w/o LLM cards	7.70×10^{-5}	3.13×10^{-5}	14.2	0.291
w/o LLM warm start	6.77×10^{-5}	2.54×10^{-5}	14.6	0.293
w/o RF frontier	2.16×10^{-5}	1.95×10^{-6}	17.4	0.261
w/o bottleneck	7.93×10^{-5}	3.87×10^{-5}	19.2	0.284
Fixed gate	6.65×10^{-5}	4.34×10^{-5}	18.8	0.254
w/o diversity	5.69×10^{-5}	3.45×10^{-5}	22.0	0.278

E.2 MODEL GROUPS

Table 7: LLM model sensitivity on the five retained model folders. All groups use five matched seeds.

Model	Final HV mean $\pm$ std	First hit mean $\pm$ std
deepseek-v4-pro	$9.23 \times 10^{-5} \pm 4.15 \times 10^{-5}$	12.2 ± 6.9
GLM-5.3	$6.09 \times 10^{-5} \pm 2.97 \times 10^{-5}$	25.6 ± 20.5
Kimi-K3	$7.64 \times 10^{-5} \pm 4.96 \times 10^{-5}$	28.2 ± 10.3
MiniMax-m3	$7.50 \times 10^{-5} \pm 4.29 \times 10^{-5}$	22.4 ± 15.2
Qwen3.6	$7.69 \times 10^{-5} \pm 3.37 \times 10^{-5}$	23.4 ± 18.0

F PROMPT CONTRACT AND AUDIT TRAIL

The online prompt is deliberately compact and structured. A representative payload is:

```
Task: five-objective molecular candidate selection
(ablation protocol).
Objectives: electronic_gap maximize at fixed threshold;
dipole_moment minimize at fixed threshold; ...
Witness id: W2. Normalized margins: [-0.18, 0.31, -0.04,
...].
Recent evidence (at most 20 completed points): each record
contains candidate id, normalized inputs, real outputs,
source (warm-start/card/safe/model), predicted uncertainty,
observed margin change, Pareto status, and incremental HV.
Previous card audit: card W2-03 had adherence 0.62 and
positive observed reward; card W2-01 was revoked after
repeated negative reward.
Return JSON cards only: card id, witness id, variable
directions, improve objectives, protect objectives, local
condition, confidence, falsification test, and abstain flag.
Do not return candidate ids, hidden labels, target values,
or GP pseudo-observations.
```

The response is accepted only if it passes schema validation, has a single witness owner, uses known variables and legal directions, and contains no hidden candidate labels. Failed calls are retried and logged; accepted cards are scored against later real observations. This contract makes the LLM contribution an explicit and auditable decision-layer component.